

BELLMAN FUNCTION TECHNIQUE IN HARMONIC ANALYSIS

ALEXANDER VOLBERG

ABSTRACT. It is strange but fruitful to think about the functions as random processes. Any function can be viewed as a martingale (in many different ways) with discrete time. But it can be useful to have continuous time too. Processes can emulate functions, expectation of profit functional on the solution of stochastic differential equation can emulate the functional on usual familiar functions. The advantage is that now we have “all” admissible functions “enumerated” as solutions of one stochastic differential equation, and choosing the best function optimizing a given functional becomes a problem of choosing the right control process. But such problem has been long since solved in the part of mathematics called Stochastic Optimal Control. So-called Bellman equation reduces an infinite dimensional problem of choosing the best control to a finite dimensional (but non-linear as a rule) PDE called Bellman equation. Its solution, called Bellman function of a given optimization problem, gives us a lot of information about optimum and optimizers. This method gave some interesting results in the classical Harmonic Analysis, having on the surface nothing to do with probability. Sometimes the results obtained by this method did not find “classical” proofs so far. It is especially well-suited to estimates of singular integrals, probably because of the underlying probabilistic structure of classical singular integrals.

1. QUASICONFORMAL MAPS: SHARP DISTORTION ESTIMATES AND SHARP REGULARITY

We deal first with Beltrami equation

$$(1) \quad f_{\bar{z}} - \mu f_z = 0,$$

with bounded function μ called *Beltrami coefficient*, for simplicity μ is compactly supported on \mathbb{C} , f being analytic near ∞ (see (1)) supposed to have the following Laurent decomposition at infinity

$$(2) \quad f(z) = z + c_0 + \frac{c_{-1}}{z} + \dots$$

If μ is smooth it is not difficult to see that the solution is smooth on the whole \mathbb{C} . But we are interested in just measurable bounded μ :

$$(3) \quad \|\mu\|_{L^\infty(\mathbb{C})} = k < 1.$$

Several natural questions appear:

1. What is the smoothness of f depending on μ, k ?
2. What are distortion properties of f ? How it distorts the area and other measures?
3. In what classes (Sobolev, say) we can solve (1) in such a way that it will be continuous $\hat{\mathbb{C}} \rightarrow \hat{\mathbb{C}}$, where $\hat{\mathbb{C}} = \mathbb{C} \cup \infty$? As at infinity it is a perfect holomorphic map, this question concerns only finite part of \mathbb{C} , so it is local.

Denote $g := f_{\bar{z}}$. It is a function with compact support. We can restore f :

$$(4) \quad f(z) = \frac{1}{\pi} \int_{\mathbb{C}} \frac{1}{\zeta - z} g(\zeta) dm_2(\zeta) + c_0 + z.$$

We used (2) and we naturally assume integrability of g .

Then obviously

$$f_z = \frac{1}{\pi} \int_{\mathbb{C}} \frac{1}{(\zeta - z)^2} g(\zeta) dm_2(\zeta) + 1,$$

where the integral is singular and should be understood, e.g. in the sense of principal values. This is an important operator called the Ahlfors–Beurling transform (AB transform):

$$Tg := \frac{1}{\pi} \int_{\mathbb{C}} \frac{1}{(\zeta - z)^2} g(\zeta) dm_2(\zeta).$$

Then (1) automatically becomes

$$(5) \quad g - \mu Tg = (I - \mu T)g = h,$$

where $h = \mu$ is bounded with compact support. So in particular $h \in \cap_{p \geq 1} L^p(\mathbb{C})$.

It is easy to make Fourier analysis of convolution kernel $\frac{\pi}{z^2}$ of AB operator, and to see that it is the Fourier multiplier with symbol $\zeta/\bar{\zeta}$. Therefore, $\|T\|_{L^2(\mathbb{C})} = 1$ and having then $\|\mu T\|_{L^2(\mathbb{C})} \leq k < 1$, we can conclude that (5) has a solution in $L^2(\mathbb{C})$ given by the usual Neumann series:

$$(6) \quad g = h + \mu Th + \mu T\mu Th + \dots,$$

Notice that g is compactly supported (as μ is). Restore f by (4). The boundedness of T in L^2 implies now that

$$(7) \quad f \in W_{1,loc}^2(\mathbb{C}).$$

Let us see now that g given in (6) is actually better than in L^2 . Operator T has norm 1 in L^2 and it has norm *close* to 1 in L^p , $p > 2$, $p \approx 2$. In fact, it is an operator with Calderón–Zygmund kernel, and as such it is bounded in all L^p . Interpolating between, say L^2 and L^4 , we get

$$(8) \quad \|T\|_{L^p(\mathbb{C})} =: n(p) \rightarrow 1, \quad p \rightarrow 2.$$

So we can find such a $p = p(k) = 2 + \varepsilon(k)$, $\varepsilon(k) > 0$, that the series in (6) converges in this $L^{2+\varepsilon(k)}$. So $g \in L^{2+\varepsilon(k)}$. Again restore f by formula (4) (it is the same f of course), again use that $f_z = Tf_{\bar{z}} + 1 = Tg + 1$, and that T is bounded in all L^p being a Calderón–Zygmund operator. We got that f self-improves from (7) to

$$(9) \quad f \in W_{1,loc}^p(\mathbb{C}), \quad p = 2 + \varepsilon(k), \quad \varepsilon(k) > 0.$$

We formulate this small fact as a fundamental Ahlfors–Bers–Bojarski’s theorem:

Theorem 1.1. *Any solution of (1) in $W_{1,loc}^2$ self-improves to being in $W_{1,loc}^{2+\varepsilon(k)}$, $\varepsilon(k) > 0$. In particular, any such solution is continuous on \mathbb{C} and even Hölder continuous. There exists a solution, which is a homeomorphism of $\hat{\mathbb{C}}$ into itself.*

New questions appear:

4. What is the largest $2 + \varepsilon(k)$?
5. What is $n(p)$ in (8)?

To this we want to add some more questions. Introduce a constant

$$K = \frac{1+k}{1-k} \in [1, \infty); \quad k = \frac{K-1}{K+1}.$$

It has a geometric meaning: it gives the maximal ratio of the axis of infinitesimal ellipses obtained as the images of infinitesimal circles by all possible solutions of (1).

Definition. Any solution of (1) from Theorem 1.1 is called K-quasiregular map. Any homeomorphic solution is called K-quasiconformal map or K-quasiconformal homeomorphism. It is basically unique because by normalization at infinity it can be only shifted.

Again questions:

6. What is the sharp distortion of K-quasiconformal maps? Namely, if $f(\mathbb{D}) = \mathbb{D}$, where \mathbb{D} denotes the unit disc and $f(0) = 0$, then what is the *sharp* (largest) exponent in

$$(10) \quad \forall E \subset \mathbb{D}, \quad |f(E)| \leq C_K |E|^{e(K)}?$$

Without normalizations, allowing f to be any K-quasiconformal map this becomes the question what is the best (largest) exponent in

$$(11) \quad \forall E \subset B, \quad \frac{|f(E)|}{|f(B)|} \leq C_K \left(\frac{|E|}{|B|} \right)^{e(K)}?$$

Function

$$f_0(z) := z|z|^{\frac{1}{K}-1}, \quad |z| \leq 1, \quad \text{and} \quad = z, \quad \text{for } |z| > 1$$

shows that $e(K) \leq \frac{1}{K}$.

Gehring's problem: $e(K) = \frac{1}{K}$. It is equivalent to saying (we will see that) in Question 4 the sharp exponent of Sobolev integrability is $2 + \varepsilon(k) = 1 + \frac{1}{k}-$. This is very tough, but it was done by Astala [1].

Glance now at (6): it gives that if we want to show that the exponent p of Sobolev integrability goes up to $1 + \frac{1}{k}-$, it is enough to prove that

$$\|T\|_{L^{1+1/k}} = 1/k,$$

in other words that

$$(12) \quad \|T\|_{L^p} = p-1, \quad p > 2.$$

This is very open, we will show how Bellman function gives partial results.

Big Iwaniec's problem or $p-1$ -problem: $n(p) = \max(p-1, \frac{p}{p-1}-1)$.

Yet another question naturally arises: in Theorem 1.1 we started with a priori solution in $W_{1,loc}^2$. How much *below* this we can start to have the same self-improvement?

7. Let $f \in W_{1,loc}^q$ solves (1), and $q \in (1, 2)$. What is the smallest $q = q(k)$ such that we still have for each such f self-improvement to $W_{1,loc}^2$? (And then automatically to $W_{1,loc}^{2+\varepsilon(k)}$ by Theorem 1.1, and then up to $W_{1,loc}^{1+1/k-}$ by Astala's [1]?)

Iwaniec's problem: $q(K) = 1 + k$.

We will prove it here using Bellman function technique and Astala's sharp distortion result [1]. We follow the exposition of [2].

1.1. Invertibility of Beltrami operator. If Big Iwaniec's problem were solved than we would immediately get

$$(13) \quad \text{If } p \in [2, 1 + 1/k), \text{ then } \|(I - \mu T)^{-1}\|_{L^p} \leq \frac{C(k)}{1 + \frac{1}{k} - p}.$$

(Actually even with $C(k) = 1/k$, but this we do not care about as k is fixed and we vary p .)

By duality and small talk one would get

$$(14) \quad \text{If } p \in I_k := (1 + k, 1 + 1/k), \text{ then } \|(I - \mu T)^{-1}\|_{L^p} \leq \frac{C(k)}{\text{dist}(p, \mathbb{R} \setminus I_k)}.$$

Big Iwaniec's conjecture is still a conjecture, but this is a Theorem of Petermichl-Volberg [58], which we start to prove now. It will use Bellman function technique. Notice that now there exists an even more precise version of this result, namely, see [3].

Theorem 1.2. *If $p \in I_k := (1 + k, 1 + 1/k)$, then $\|(I - \mu T)^{-1}\|_{L^p} \leq \frac{C(k)}{\text{dist}(p, \mathbb{R} \setminus I_k)}$.*

Let f be a K -quasiconformal homeomorphism, let $p \in [2, 1 + 1/k)$, where $k = \frac{K-1}{K+1}$. Denote $J_f = |f_z|^2 - |f_{\bar{z}}|^2$, the Jacobian of f . We need lemma:

Lemma 1.3. *Let f, p be as above, denote $w := J_f^{1-p/2}$. Then $w \in A_2$ with*

$$[w]_{A_2} \leq \frac{p^2 C(K)}{1 + \frac{1}{k} - p}.$$

Remark. We will give the proof following [2]. There is another very interesting proof in [3].

Proof. Notice first that

$$(15) \quad (1 - k^2)|f_z|^2 \leq J_f = |f_z|^2 - |f_{\bar{z}}|^2 \leq |f_z|^2 \leq |f_z|^2 + |f_{\bar{z}}|^2.$$

That is all this quantities are comparable with $C = C(K)$. The next step is to show that there is $C(K)$ such that if $B \subset \mathbb{C}$ is a disc and if f is a K -quasiconformal homeomorphism of \mathbb{C} , then

$$(16) \quad \frac{1}{|B|} \int_B (|f_z| + |f_{\bar{z}}|)^p \leq \frac{pC(K)}{1 + \frac{1}{k} - p} \left(\frac{|f(B)|}{|B|} \right)^{\frac{p}{2}}.$$

Using linear maps to pre-compose and to post-compose with f we reduce it to normalized case $|f(B)| = |B| = 1$. Apply (11) proved by Astala in [1] to the set

$$E_t = \{z \in B : |f_z|^2 + |f_{\bar{z}}|^2 \geq t\}, \quad t > 0$$

we get

$$\begin{aligned} |E_t| &\leq \frac{1}{t} \int_{E_t} (|f_z|^2 + |f_{\bar{z}}|^2) dm_2 \leq K \frac{1}{t} \int_{E_t} (|f_z|^2 - |f_{\bar{z}}|^2) dm_2 = \\ &K \frac{1}{t} |f(E_t)| \leq C_1(K) \frac{1}{t} |E_t|^{\frac{1}{K}}. \end{aligned}$$

Therefore,

$$|E_t| \leq \min(1, C_2(K) \frac{1}{t^{\frac{K}{K-1}}}).$$

This is the same as

$$|\{z \in B : |f_z| + |f_{\bar{z}}| \geq t\}| \leq \min(1, C_3(K) \frac{1}{t^{\frac{2K}{K-1}}}).$$

Distribution function calculation now shows

$$(17) \quad \int_B (|f_z| + |f_{\bar{z}}|)^p \leq C' + C'' p \int_1^\infty \frac{t^{p-1}}{t^{\frac{2K}{K-1}}} = C' + C'' p \int_1^\infty \frac{1}{t^{2+\frac{1}{k}-p}} dt \leq C' + C'' p \frac{1}{1+\frac{1}{k}-p},$$

as $\frac{2K}{K-1} = 1 + \frac{1}{k}$. This proves (16).

Now we are ready to prove Lemma 1.3. Notice that $w = J_f^{1-p/2} = (J_{f^{-1}} \circ f)^{p/2-1}$. Then

$$\frac{1}{|B|} \int w \, dm_2 = \frac{1}{|B|} \int_B (J_{f^{-1}} \circ f)^{p/2-1}(z) \, dm_2(z) = \frac{1}{|B|} \int_{f(B)} J_{f^{-1}}^{p/2}(\zeta) \frac{J_{f^{-1}}(\zeta)}{J_{f^{-1}}(\zeta)} \, dm_2(\zeta),$$

where we made the change of variable $z = f^{-1}(\zeta)$. We continue

$$\frac{1}{|B|} \int_B w \, dm_2 = \frac{|f(B)|}{|B|} \frac{1}{|f(B)|} \int_{f(B)} J_{f^{-1}}^{p/2}(\zeta) \, dm_2(\zeta) \leq \frac{pC(K)}{1+\frac{1}{k}-p} \left(\frac{|B|}{|f(B)|} \right)^{p/2} \frac{|f(B)|}{|B|}.$$

So

$$(18) \quad \frac{1}{|B|} \int_B w \, dm_2 \leq \frac{pC(K)}{1+\frac{1}{k}-p} \left(\frac{|B|}{|f(B)|} \right)^{p/2-1}.$$

We used here (16) with K -**quasidisc** $f(B)$ instead of a disc B . But this does not matter as any K -*quasidisc* ($:=$ the image of a disc by K -quasiconformal map) is an almost disc with constants depending only on K .

Now notice that we assumed $p \geq 2$, so if $p_n := p - 2$ we can write

$$\frac{1}{|B|} \int_B w^{-1} \, dm_2 = \frac{1}{|B|} \int_B (J_f)^{p/2-1}(z) \, dm_2(z) = \frac{1}{|B|} \int_B J_f^{p_n/2}(z) \, dm_2(z),$$

and we use again (16) with p_n replacing p , gives us:

$$(19) \quad \frac{1}{|B|} \int_B w^{-1} \, dm_2 \leq \frac{p_n C(K)}{1+\frac{1}{k}-p_n} \left(\frac{|f(B)|}{|B|} \right)^{p_n/2} = \frac{\max(C', p-2)C(K)}{3+\frac{1}{k}-p} \left(\frac{|f(B)|}{|B|} \right)^{p/2-1}.$$

□

Multiplying (18) and (19) we get Lemma 1.3:

$$(20) \quad w = J_f^{1-\frac{p}{2}}, p \in [2, 1 + \frac{1}{k}) \Rightarrow [w]_{A_2} \leq \frac{p^2 C(K)}{1 + \frac{1}{k} - p}.$$

Now we can reap a first consequence:

Theorem 1.4. *Suppose $\|\mu\|_\infty = k < 1$, let $p \in I_k := (1 + k, 1 + \frac{1}{k})$. Then operators $I - \mu T, I - T\mu$ are boundedly invertible in L^p .*

Proof. We can work with $I - \mu T$ as $I - T\mu = T(I - \mu T)T^{-1}$ and T is boundedly invertible in each L^q , $1 < q < \infty$ (T^{-1} is again a Fourier multiplier of Calderón–Zygmund type).

Suppose we know how to prove the estimate from below

$$(21) \quad \|(I - \mu T)g\|_p \geq c(p, k)\|g\|_p, \forall g \in L^p(\mathbb{C}), p \in I_k.$$

Then we would now exactly the same for $I - \mu T$, and, so, for the adjoint operator $(I - \mu T)^*$. Then $I - \mu T$ would have dense images in all L^p we consider. Joining this with the estimate from below (21) we would conclude that $I - \mu T$ are invertible in all L^p , $p \in I_k$.

So it is enough to have (21). And we would like a good estimate for $c(K, p)$ in it.

It is enough to prove (21) for the dense set of functions

$$g \in C_0^\infty(\mathbb{C}), \int_{\mathbb{C}} g \, dm_2 = 0.$$

Let ϕ be the Cauchy transform of g : $\phi = \frac{1}{\pi} \int \frac{g(\zeta)}{\zeta - z} \, dm_2(\zeta)$.

Denoting $h := g - \mu Tg$ we come to equation

$$\phi_{\bar{z}} - \mu \phi_z = h, \text{ in which we want to estimate } \|\phi_{\bar{z}}\|_p \leq C(K, p)\|h\|_p \text{ if } p \in I_k.$$

By Theorem 1.1 there is a K -qc homeomorphism f satisfying $f_{\bar{z}} - \mu f_z = 0$. Set

$$u = \phi \circ f^{-1},$$

and let us see how equation $\phi_{\bar{z}} - \mu \phi_z = h$ will be transformed by this change of variable.

We calculate

$$\begin{aligned} \phi_{\bar{z}} &= (u_z \circ f)f_{\bar{z}} + (u_{\bar{z}} \circ f)\bar{f}_z, \\ \phi_z &= (u_z \circ f)f_z + (u_{\bar{z}} \circ f)\bar{f}_{\bar{z}}, \\ \phi_{\bar{z}} - \mu \phi_z - h &= (u_z \circ f)f_{\bar{z}} + (u_{\bar{z}} \circ f)\bar{f}_z - \mu((u_z \circ f)f_z + (u_{\bar{z}} \circ f)\bar{f}_{\bar{z}}) - h = \\ &= (u_z \circ f)\bar{f}_z - \mu(u_{\bar{z}} \circ f)\bar{\mu}\bar{f}_z - h = (1 - |\mu|^2)(u_{\bar{z}} \circ f)\bar{f}_z - h = 0. \end{aligned}$$

Hence obviously

$$\int |u_{\bar{z}} \circ f|^p |f_z|^p \leq C(K) \int |h|^p \Rightarrow \int |u_{\bar{z}} \circ f|^p J_f^{p/2-1} J_f \leq C(K) \int |h|^p$$

And changing variable we get

$$(22) \quad \int |u_{\bar{z}}|^p |(J_{f^{-1}})^{1-p/2}| = \int |u_{\bar{z}}|^p |(J_f \circ f^{-1})^{p/2-1}| \leq C(K) \int |h|^p$$

On the other hand,

$$\int |u_z \circ f|^p |f_{\bar{z}}|^p \leq \frac{k^2}{1 - k^2} \int |u_z \circ f|^p J_f^{p/2-1} J_f = \frac{k^2}{1 - k^2} \int |u_z|^p (J_{f^{-1}})^{1-p/2}$$

Denote by $W := (J_{f^{-1}})^{1-p/2}$. It is the one in Lemma 1.3, only f replaced by f^{-1} , which is a K -qc homeomorphism as well.

But the last expression above can be written as

$$\int |u_z|^p (J_{f^{-1}})^{1-p/2} = \int |T(u_{\bar{z}})|^p (J_{f^{-1}})^{1-p/2} = \int |T(u_{\bar{z}})|^p W =: TU$$

(in fact, u can be restored from $u_{\bar{z}}$ by Cauchy integral with no addition because u vanishes at infinity; then $u_z = T(u_{\bar{z}})$). Notice that (22) says that

$$(23) \quad \int |u_{\bar{z}}|^p W \leq C(K) |h|^p.$$

Combine (23) with (here F is some “unknown” function on $[1, \infty)$, but finite for all finite arguments)

$$TU = \int |T(u_{\bar{z}})|^p W = F([w]_{A_2}) \int |u_{\bar{z}}|^p W$$

to get

$$(24) \quad \|\phi_{\bar{z}}\|_p \leq C \left(\int |u_{\bar{z}}|^p W + \int |T(u_{\bar{z}})|^p W \right) \leq C(K) (1 + F([w]_{A_2})) \|h\|_p^p.$$

Noticing that Lemma 1.3 gives the estimate $[w]_{A_2} \leq \frac{p^2 C(K)}{1 + \frac{1}{k} - p}$, we conclude finally that

$$\|g\|_p \leq C(K) F\left(\frac{p^2 C(K)}{1 + \frac{1}{k} - p}\right) \|h\|_p, \text{ if } p \in [2, 1 + \frac{1}{k}).$$

We need the same estimate now for $1 + k < p \leq 2$. We need $W \in A_p$ (now $p \leq 2$). But it is the same as to say that $W^{-1/(p-1)} \in A_{p'}, p' = p/(p-1)$. In our case $W := (J_{f-1})^{1-p/2}$, so $W^{-1/(p-1)}$ will be $(J_{f-1})^{\frac{p'}{2}-1}$, which is inverse to the one in Lemma 1.3, so also in $A_2 \subset A_{p'}$. We get for the whole interval of p 's: $p \in I_k = (1 + k, 1 + \frac{1}{k})$ also

$$(25) \quad \|g\|_p \leq F\left(\max\left(\frac{p^2 C(K)}{1 + \frac{1}{k} - p}, \frac{p^2 C(K)}{1 + \frac{1}{k} - p'}\right)\right) \|h\|_p = F\left(\frac{p^2 C(K)}{\text{dist}(p, \mathbb{R} \setminus I_k)}\right) \|h\|_p.$$

Theorem 1.1 is proved. □

In [2] the following conjecture was formulated that claims that function F in (25) is just linear. Notice that this would in fact easily follow from Big Iwaniec' conjecture.

Conjecture

$$(26) \quad \|g\|_p \leq \frac{p^2 C(K)}{\text{dist}(p, \mathbb{R} \setminus I_k)} \|h\|_p, \text{ equivalently } \|(I - \mu T)^{-1}\|_p \leq \frac{p^2 C(K)}{\text{dist}(p, \mathbb{R} \setminus I_k)}.$$

We will prove now this conjecture using the Bellman function technique. But first let us derive the corollary of the conjecture. As always $\|\mu\|_\infty = k < 1$.

Theorem 1.5 (Corollary of the conjecture). *Any solution of*

$$F_{\bar{z}} - \mu F_z = 0,$$

which is in $W_{1,loc}^{1+k}$ is automatically in $W_{1,loc}^2$, and so satisfies Theorem 1.1. It automatically self-improves then (by Astala's [1]) to be in $W_{1,loc}^{1+\frac{1}{k}-}$.

First use Conjecture to prove

Lemma 1.6 (Behavior at the end points of interval I_k). *Operators $I - \mu T, I - T\mu$ have dense range in $L^{1+\frac{1}{k}}$ and, correspondingly, trivial kernels on L^{1+k} .*

Proof. By $T(I - \mu T)T^{-1} = I - T\mu$ and invertibility of T in all spaces L^p , $1 < p < \infty$, it is enough to prove just the dense range of $I - \mu T$ in $L^{1+\frac{1}{k}}$. Consider $\varepsilon > 0$ and equation

$$(27) \quad \phi_\varepsilon - (1 - \varepsilon)\mu T\phi_\varepsilon = h$$

for nice $h \in C_0^\infty$. We want to consider the solution for $p_0 = 1 + \frac{1}{k}$. We consider this p_0 in $I_{(1-\varepsilon)k}$ because $\|(1 - \varepsilon)\mu\|_\infty = (1 - \varepsilon)k$. Point p_0 is obviously $C(K)\varepsilon$ close to the right end point of $I_{(1-\varepsilon)k}$.

Hence, applying conjecture we conclude that

$$\|\phi_\varepsilon\|_{p_0} \leq \frac{C(K)}{\varepsilon} \|h\|_{p_0}.$$

Notice two things: 1) In L^2 the norma of ϕ_ε are uniformly bounded by $C(K)$ just by using Neumann series in L^2 in (27); 2) in L^{p_0} the norms of $T(\varepsilon\phi_\varepsilon)$ are uniformly bounded. It is immediate to conclude from 1) and 2) that

$$\varepsilon\mu T\phi_\varepsilon \text{ converges weakly to zero in } L^{p_0}.$$

Rewrite our equation (27) as follows:

$$\phi_\varepsilon - \mu T\phi_\varepsilon = h - \varepsilon\mu T\phi_\varepsilon.$$

The right hand side weakly in L^{p_0} converges to any function h , whose family is strongly dense in L^{p_0} . So the right hand side is weakly dense in $L^{p_0} = L^{1+\frac{1}{k}}$. But it is in $\text{Range}(I - \mu T)$, so this range is weakly dense in $L^{1+\frac{1}{k}}$. Being a linear set this range is then strongly dense in $L^{1+\frac{1}{k}}$. Lemma 1.6 is proved □

The proof of Theorem 1.5. Consider $R_{\bar{z}} - \mu R_z = 0$, $R \in W_{1,loc}^{1+k}$. Choose $\phi \in C_0^\infty$. Set $G = \phi R$. Then

$$G_{\bar{z}} - G_z = (\phi_{\bar{z}} - \mu\phi_z)R.$$

Looking at this formula we can start to think that the support of μ is compact (is contained in the support of ϕ).

As G vanishes at infinity it is the Cauchy transform of its $\bar{\partial}G = G_{\bar{z}}$, and therefore, $G_z = TG_{\bar{z}}$. We can rewrite the equation

$$(I - \mu T)\psi = h; \quad \psi := G_{\bar{z}}, \quad h = (\phi_{\bar{z}} - \mu\phi_z)R \in L^{\frac{2(1+k)}{1-k}} \subset L^2(\mathbb{C}) \cap L^{2+\varepsilon}(\mathbb{C}).$$

The inclusion above for function R is by Sobolev imbedding, in fact, we assumed that $R \in W_{1,loc}^{1+k}$, and by the compactness of the support of R . It has been already remarked, that in the last two equation we have the right to think that $\mu = 0$ outside of the support of ϕ . Let us consider the convergent in $L^2(\mathbb{C})$ of the series of compactly supported functions:

$$\psi_0 = h + \mu Th + \mu T\mu Th + \dots$$

It solves our equation, it is in $L^2(\mathbb{C})$ and it is compactly supported, hence it is in $L^{1+k}(\mathbb{C})$. Therefore we got a solution ψ_0 of $(I - \mu T)\psi_0 = h$, which is in $L^{1+k} \cap L^2$. But $\psi = G_{\bar{z}}$ is also in L^{1+k} . By Lemma 1.6 we have $\psi = \psi_0 \in L^2$. It means that $R \in W_{1,loc}^2$. Theorem 1.5 is proved. □

Example. $f = \frac{|z|^{1-\frac{1}{K}}}{z}$ for $z \in \mathbb{D}$, $f = \frac{1}{z}$ outside \mathbb{D} . It is a solution of Beltrami equation with $\mu, \|\mu\|_\infty = k = \frac{K-1}{K+1}$, and it is in $W_{1,loc}^q$ for every $q < 1+k$. But it is NOT K -quasiregular mapping, it has a singularity at 0.

This example shows how sharp is Theorem 1.5. Its proof hinges on Conjecture 26. We prove this Conjecture now using Bellman technique. First we analyze function F from (25). Recall that $W := (J_{f^{-1}})^{1-p/2}$, and if $p \in (1+k, 1+\frac{1}{k})$ the estimate in (25) is $F_p([W]_{A_2})$, where F_p is the best function one can have in the estimate

$$\|T\|_{L^p(W)} \leq F_p([W]_{A_p}).$$

Suppose we can prove

Theorem 1.7. $F(x) \leq C x^{\max(1, 1/(p-1))}$.

Then we recall that for $1+k < p \leq 2$, $p' := \frac{p}{p-1}$ we already estimated $[W]_{A_p}^{1/(p-1)} = [W^{-1/(p-1)}]_{A_{p'}} \leq [W^{-1/(p-1)}]_{A_2} \leq \frac{p^2 C(K)}{p-1-k}$. And for $1+\frac{1}{k} > p \geq 2$, $[W]_{A_2} \leq \frac{p^2 C(K)}{1+\frac{1}{k}-p}$. These estimates were based on sharp distortion theorem of Astala [1]. We made these estimates in Lemma 1.3. These estimates and Theorem 1.7 then imply trivially conjecture (26). Hence this Theorem 1.7 is the only ingredient left to be proved to have Theorem 1.5.

2. LINEAR ESTIMATES OF WEIGHTED AHLFORS–BEURLING TRANSFORM BY BELLMAN FUNCTION TECHNIQUE

Let ω be any weight on \mathbb{R}^2 , denote its heat extension into \mathbb{R}_+^3 by $\omega(x, t) = \omega(x_1, x_2, t)$:

$$\omega(x, t) = \frac{1}{\pi t} \int \int_{\mathbb{R}^2} \omega(y) \exp\left(-\frac{\|x-y\|^2}{t}\right) dy_1 dy_2.$$

We define

$$[\omega]_{A_p}^{heat} := \sup_{(x,t) \in \mathbb{R}_+^3} \omega(x, t) \left(\omega^{-\frac{1}{p-1}}(x, t) \right)^{p-1}.$$

The weights w with finite $[w]_{A_p}^{heat}$ are called A_p weights. There is an extensive theory of A_p weights, see for example [55], [40]. The usual definition differs from the one above, but it describes the same class of weights. Actually, we will say more about the relationship between the classical definition and ours. But first we state two more theorems, whose combined use gives Theorem 1.7 at least for $p \geq 2$.

Remark. The method called Rubio de Francia extrapolation—one can see its exposition in [27]—actually shows that to have a full range of p 's in Theorem 1.7 it is enough to prove it only for $p = 2$.

Theorem 2.1. *For any A_p weight w and any $p \geq 2$ we have*

$$\|T\|_{L^p(wdA) \rightarrow L^p(wdA)} \leq C(p) ([w]_{A_p}^{heat})^{\frac{1}{p-1}}.$$

We want to discuss the connection between $[w]_{A_p}^{heat}$ and $[w]_{A_p}^{class}$. Here $[w]_{A_p}^{class}$ denotes the following supremum over all discs in the plane:

$$[w]_{A_p}^{class} := \sup_{B(x,R)} \left(\frac{1}{|B(x,R)|} \int_{B(x,R)} \omega dA \right) \cdot \left(\frac{1}{|B(x,R)|} \int_{B(x,R)} \omega^{-\frac{1}{p-1}} dA \right)^{p-1}.$$

Obviously, there exists a positive absolute constant a such that for any function w

$$a [w]_{A_p}^{class} \leq [w]_{A_p}^{heat}.$$

Remark. The opposite inequality is easy to prove too. In fact, we have

Theorem 2.2. *There exists a finite absolute constant b such that*

$$[w]_{A_p}^{heat} \leq b [w]_{A_p}^{class}.$$

Proof. Constants will be denoted by the letters c, C ; they may vary from line to line and even within the same line. We introduce the following notations. B_k denotes $B(0, 2^k)$, $k = 0, 1, 2, \dots$, $\langle f \rangle_B$ stands for the average $\frac{1}{|B|} \int_B f dA$, $f(B)$ stands for $\int_B f dA$. If $B = B(0, r)$, then $\langle f \rangle_B^h$ stands for $\frac{1}{\pi r^2} \iint_{\mathbb{R}^2} f(x) \exp(-\frac{\|x\|^2}{r^2}) dx_1 dx_2$.

Lemma 2.3. *Suppose f and g , positive functions on the plane, are such that $\sup_B \langle f \rangle_B \langle g \rangle_B = A$, then there exists a finite absolute constant c such that*

$$\langle f \rangle_B \langle g \rangle_B^h \leq cA$$

for any disc B .

Proof. Scale invariance allows us to prove this only for one disc $B = B(0, 1)$. We start the estimate:

$$\langle f \rangle_B \langle g \rangle_B^h \leq c \langle f \rangle_{B_{\Sigma_k}} 2^{2k} \exp(-2^{2k-2}) \frac{A}{\langle f \rangle_{B_k}}.$$

On the other hand $\langle f \rangle_{B_k} > c \langle f \rangle_{B_{k-1}} > \dots c^k \langle f \rangle_B$ (recall that B is the unit disc). Plugging this in the inequality above, we get

$$\langle f \rangle_B \langle g \rangle_B^h \leq c \langle f \rangle_{B_{\Sigma_k}} C^k \exp(-2^{2k-2}) \frac{A}{\langle f \rangle_B}.$$

In other words,

$$\langle f \rangle_B \langle g \rangle_B^h \leq c A \Sigma_k C^k \exp(-2^{2k-2}) = cA$$

and the lemma is proved. \square

Now we want to prove Theorem 2.2. Fix B . Again by scale invariance it is enough to consider $B = B(0, 1)$. By the previous lemma, we know that

$$(28) \quad \langle f \rangle_{B_k} \langle g \rangle_{B_k}^h \leq cA$$

for any k .

Now

$$\langle f \rangle_B^h \langle g \rangle_B^h \leq c \langle g \rangle_B^h \Sigma 2^{2k} \exp(-2^{2k-2}) \langle f \rangle_{B_k} \leq c \langle g \rangle_B^h \Sigma 2^{2k} \exp(-2^{2k-2}) \frac{cA}{\langle g \rangle_{B_k}^h}.$$

The last inequality used (28).

On the other hand, $\langle g \rangle_{B_k}^h > c \langle g \rangle_{B_{k-1}}^h > \dots c^k \langle g \rangle_B^h$ (recall that B is the unit disc). Plugging this in the inequality above, we get

$$\langle f \rangle_B^h \langle g \rangle_B^h \leq c \langle g \rangle_B^h \Sigma_k C^k \exp(-2^{2k-2}) \frac{cA}{\langle g \rangle_B^h}.$$

In other words,

$$\langle f \rangle_B^h \langle g \rangle_B^h \leq c^2 A \Sigma_k C^k \exp(-2^{2k-2}) = c^2 A.$$

Theorem 2.2 is completely proved. \square

The next result proves Theorem 1.7 for $p = 2$. We will show later how to extrapolate just from the result at $p = 2$ to all possible p 's.

Theorem 2.4. *For any A_2 weight w we have*

$$\|T\|_{L^2(wdA) \rightarrow L^2(wdA)} \leq C [w]_{A_2}^{class}.$$

Proof. There will be many steps. But we are going to prove only Theorem 2.1 and only for $p = 2$. By Theorem 2.2 and Rubio de Francia extrapolation this is enough.

The operator T is given in the Fourier domain (ξ_1, ξ_2) by the multiplier $\frac{\xi}{|\xi|^2} = \frac{(\xi_1 + i\xi_2)^2}{\xi_1^2 + \xi_2^2} = \frac{\xi_1^2}{\xi_1^2 + \xi_2^2} - \frac{\xi_2^2}{\xi_1^2 + \xi_2^2} + 2i \frac{\xi_1 \xi_2}{\xi_1^2 + \xi_2^2}$. Thus, T can be written as $T = R_1^2 - R_2^2 + 2iR_1R_2$, where R_1, R_2 are Riesz transforms on the plane (see [55] for their definition and properties). Another way of writing T is

$$T = m_1 + im_2,$$

where m_1, m_2 are Fourier multiplier operators. Notice that the multipliers themselves (as functions, not as multiplier operators) are connected by

$$m_2 = m_1 \circ \rho,$$

where ρ is $\pi/4$ rotation of the plane. So the multiplier operators are related by

$$m_2 = U_\rho m_1 U_\rho^{-1},$$

where U_ρ is an operator of ρ -rotation in (x_1, x_2) plane. But for any operator K we have

$$\|U_\rho K U_\rho^{-1}\|_{L^2(wdA) \rightarrow L^2(wdA)} = \|K\|_{L^2(w \circ \rho^{-1} dA) \rightarrow L^2(w \circ \rho^{-1} dA)}.$$

Combining this with the fact that $Q_{w,2}^{heat} = Q_{w \circ \rho^{-1},2}^{heat}$ for any rotation, we conclude that we only need the desired estimate of Theorem 2.4 for $m_1 = R_1^2 - R_2^2$. Actually, we will show that

$$(29) \quad \|R_i^2\|_{L^2(wdA) \rightarrow L^2(wdA)} \leq C Q_{w,2}^{heat}, \quad i = 1, 2.$$

To prove (2.6) we fix, say, R_1^2 and two test functions $\varphi, \psi \in C_0^\infty$. We will be using heat extensions. For f on the plane, its heat extension is given by the formula

$$f(y, t) := \frac{1}{\pi t} \iint_{\mathbb{R}^2} f(x) \exp\left(-\frac{|x-y|^2}{t}\right) dx_1 dx_2, \quad (y, t) \in \mathbb{R}_+^3.$$

We usually use the same letter to denote a function and its heat extension.

Lemma 2.5. *Let $\varphi, \psi \in C_0^\infty$. Then the integral $\iiint \frac{\partial \varphi}{\partial x_1} \cdot \frac{\partial \psi}{\partial x_1} dx_1 dx_2 dt$ converges absolutely and*

$$(30) \quad \iint R_1^2 \varphi \cdot \psi dx_1 dx_2 = -2 \iiint \frac{\partial \varphi}{\partial x_1} \cdot \frac{\partial \psi}{\partial x_1} dx_1 dx_2 dt.$$

Proof. The proof of this lemma is actually trivial. It is based on the well-known fact that a function is an integral of its derivative, and also involves Parseval's formula. Consider $\varphi, \psi \in C_0^\infty$ and now

$$\begin{aligned}
\iint \psi R_1^2 \varphi dx_1 dx_2 &= \iint \frac{\xi_1^2}{\xi_1^2 + \xi_2^2} \hat{\varphi}(\xi_1, \xi_2) \hat{\psi}(-\xi_1, -\xi_2) d\xi_1 d\xi_2 = \\
&2 \iint \int_0^\infty e^{-2t(\xi_1^2 + \xi_2^2)} \xi_1^2 \hat{\varphi}(\xi_1, \xi_2) \hat{\psi}(\xi_1, \xi_2) d\xi_1 d\xi_2 dt = \\
&-2 \int_0^\infty \iint i\xi_1 \hat{\varphi}(\xi_1, \xi_2) e^{-t(\xi_1^2 + \xi_2^2)} \cdot i\xi_1 \hat{\psi}(-\xi_1, -\xi_2) e^{-t(\xi_1^2 + \xi_2^2)} d\xi_1 d\xi_2 dt = \\
&-2 \int_0^\infty \iint \frac{\partial \varphi}{\partial x_1}(x_1, x_2, t) \frac{\partial \psi}{\partial x_1}(x_1, x_2, t) dx_1 dx_2 dt = \\
&-2 \iiint_{\mathbb{R}_+^3} \frac{\partial \varphi}{\partial x_1}(x_1, x_2, t) \frac{\partial \psi}{\partial x_1}(x_1, x_2, t) dx_1 dx_2 dt.
\end{aligned}$$

Above we used Parseval's formula twice, and also we used the absolute convergence of the integrals

$$\begin{aligned}
&\iiint_{\mathbb{R}_+^3} e^{-2t(\xi_1^2 + \xi_2^2)} \xi_1^2 \hat{\varphi}(\xi_1, \xi_2) \hat{\psi}(\xi_1, \xi_2) d\xi_1 d\xi_2 dt, \\
&\iiint_{\mathbb{R}_+^3} \frac{\partial \varphi}{\partial x_1}(x_1, x_2, t) \frac{\partial \psi}{\partial x_1}(x_1, x_2, t) dx_1 dx_2 dt.
\end{aligned}$$

For the first integral this is obvious. The absolute convergence of the second integral can be easily proved. We leave this as an exercise for the reader . \square

Our next goal is to estimate the right side of (30) from above.

Theorem 2.6. *For any $\varphi, \psi \in C_0^\infty$, and any positive function w on the plane we have*

$$\iiint_{\mathbb{R}_+^3} \left| \frac{\partial \varphi}{\partial x_1} \right| \left| \frac{\partial \psi}{\partial x_1} \right| dx_1 dx_2 dt \leq A Q_{w,2}^{heat} \left(\iint |\varphi|^2 w dx_1 dx_2 + \iint |\psi|^2 \frac{1}{w} dx_1 dx_2 \right)$$

where A is an absolute constant.

Bellman function

In the proof we at last use a **Bellman function** tailored for this problem. It is B from the following theorem. The meaning of Q in the next theorem is $Q := [w]_{A_2}^{heat}$.

We use the notation H_f for the Hessian matrix of function $f : \mathbb{R}^k \rightarrow \mathbb{R}$ (the matrix of second derivatives of f), and $d^2 f$ for the second differential form, which is the quadratic form $(H_f(x)dx, dx)$, where (\cdot, \cdot) is the usual scalar product in \mathbb{R}^k , x is a point in the domain of definition of f , dx is an arbitrary vector in \mathbb{R}^k .

Theorem 2.7. *For any $Q > 1$ define the domain $D_Q := \{0 < (X, Y, x, y, r, s) : x^2 < Xs, y^2 < Yr, 1 < rs < Q\}$. Let K be any compact subset of D_Q . Then there exists a function $B = B_{Q,K}(X, Y, x, y, r, s)$ infinitely differentiable in a small neighborhood of K such that*

$$\begin{aligned} 1) \quad & 0 \leq B \leq 5Q(X + Y), \\ 2) \quad & -d^2 B \geq |dx||dy|. \end{aligned}$$

We prove Theorem 2.7 later. Now we use it to obtain the proof of Theorem 2.6.

Proof. Given a non-constant smooth w that is constant outside some large ball, we consider $Q = Q_{w,2}^{heat}$. We treat only the case $w \in A_2$, that is $Q < \infty$, for otherwise there is nothing to prove. Consider two nonnegative functions $\varphi, \psi \in C_0^\infty$. Now take $B = B_{Q,K}$, where a compact K remains to be chosen.

We are interested in

$$b(x, t) := B((\varphi^2 w)(x, t), (\psi^2 w^{-1})(x, t), \varphi(x, t), \psi(x, t), w(x, t), w^{-1}(x, t)).$$

This is a well defined function, because the choice of Q ensures that the 6-vector v , consisting of heat extensions of corresponding functions on \mathbb{R}^2 ,

$$v := ((\varphi^2 w)(x, t), (\psi^2 w^{-1})(x, t), \varphi(x, t), \psi(x, t), w(x, t), w^{-1}(x, t))$$

lies in D_Q for any $(x, t) \in \mathbb{R}_+^3$. Also we can fix any compact subset M of the open set \mathbb{R}_+^3 and guarantee that for $(x, t) \in M$, the vector v lies in a compact K . In fact, notice that for our w and for compactly supported φ, ψ the mapping $(x, t) \rightarrow v(x, t)$ maps compacts in \mathbb{R}_+^3 to compacts in D_Q . Now just take K large enough.

The main object we want to study is

$$(31) \quad \left(\frac{\partial}{\partial t} - \Delta \right) b(x, t).$$

For simplicity we assume that B is already C^2 up to the boundary of D_Q . The technical details what to do without this assumption are left to the audience, see [58]. We want to estimate the expression in (31) 1) from above *in average* and 2) from below in a pointwise way.

1) Take a “slab” $S_{\varepsilon,H} := \{(x, t) \in \mathbb{R}_+^3 : \varepsilon \leq t \leq H\}$. Notice that for any fixed positive t

$$\int_{\mathbb{R}^2} \Delta b(x, t) dx = 0.$$

This is because we assumed B to be smooth and because $v(x, t) \rightarrow 0$ for a fixed t when $x \rightarrow \infty$ rather fast, and the same is true for $\nabla v(x, t)$. Hence,

$$\int_{S_{\varepsilon,H}} \left(\frac{\partial}{\partial t} - \Delta \right) b(x, t) dx dt = \int_{S_{\varepsilon,H}} \frac{\partial}{\partial t} b(x, t) dx dt = \int_{\mathbb{R}^2} b(x, H) dx - \int_{\mathbb{R}^2} b(x, \varepsilon) dx.$$

Now we recall that $b = B \circ v$, that $B \geq 0$ (so we can throw away a “minus” term above), and that $B(X, Y, \dots) \leq 5Q(X + Y)$. Then we get (functions below are heat extensions of the corresponding symbols on \mathbb{R}^2):

$$\int_{S_{\varepsilon,H}} \left(\frac{\partial}{\partial t} - \Delta \right) b(x, t) dx dt \leq$$

$$(32) \quad 5Q \int_{\mathbb{R}^2} (\varphi^2 w(x, H) + \psi^2 w^{-1}(x, H)) dx = 5Q \int_{\mathbb{R}^2} [(\varphi^2 w)(x) + (\psi^2 w^{-1})(x)] dx.$$

2) Now we make a pointwise estimate of (31) from below. The next calculation is simple but it is key to the proof. In it as everywhere

$$v = ((\varphi^2 w)(x, t), (\psi^2 w^{-1})(x, t), \varphi(x, t), \psi(x, t), w(x, t), w^{-1}(x, t)).$$

Lemma 2.8.

$$\left(\frac{\partial}{\partial t} - \Delta \right) b(x, t) = \left((-d^2 B) \frac{\partial v}{\partial x_1}, \frac{\partial v}{\partial x_1} \right)_{\mathbb{R}^6} + \left((-d^2 B) \frac{\partial v}{\partial x_2}, \frac{\partial v}{\partial x_2} \right)_{\mathbb{R}^6}.$$

Proof.

$$\begin{aligned} \frac{\partial}{\partial t} b &= (\nabla B, \frac{\partial v}{\partial t})_{\mathbb{R}^6}, \\ \Delta b &= \left((d^2 B) \frac{\partial v}{\partial x_1}, \frac{\partial v}{\partial x_1} \right)_{\mathbb{R}^6} + \left((d^2 B) \frac{\partial v}{\partial x_2}, \frac{\partial v}{\partial x_2} \right)_{\mathbb{R}^6} + (\nabla B, \Delta v)_{\mathbb{R}^6}. \end{aligned}$$

We just used the chain rule. Now

$$\left(\frac{\partial}{\partial t} - \Delta \right) b = \left(\nabla B, \left(\frac{\partial v}{\partial t} - \Delta v \right) \right)_{\mathbb{R}^6} - \left((d^2 B) \frac{\partial v}{\partial x_1}, \frac{\partial v}{\partial x_1} \right)_{\mathbb{R}^6} - \left((d^2 B) \frac{\partial v}{\partial x_2}, \frac{\partial v}{\partial x_2} \right)_{\mathbb{R}^6}.$$

However, the first term is zero because all entries of the vector v are solutions of the heat equation. \square

By Theorem 2.7

$$-d^2 B \geq |dx||dy|.$$

Therefore, for (x, t) Lemma 2.8 gives:

$$(33) \quad \left(\frac{\partial}{\partial t} - \Delta \right) b(x, t) \geq \left| \frac{\partial \varphi}{\partial x_1} \right| \left| \frac{\partial \psi}{\partial x_1} \right| + \left| \frac{\partial \varphi}{\partial x_2} \right| \left| \frac{\partial \psi}{\partial x_2} \right|.$$

Combining (32) (33) we get

$$(34) \quad \iiint_{S_{\varepsilon, H}} \left(\left| \frac{\partial \varphi}{\partial x_1} \right| \left| \frac{\partial \psi}{\partial x_1} \right| + \left| \frac{\partial \varphi}{\partial x_2} \right| \left| \frac{\partial \psi}{\partial x_2} \right| \right) \leq 5Q \left(\iint \varphi^2 w + \iint \psi^2 w^{-1} \right).$$

Theorem 2.6 is completely proved by using a **Bellman function** of our problem whose existence is claimed in Theorem 2.7 \square

Theorem 2.4 is proved. \square

Remark. The proof of Theorem 2.6 is actually “equivalent” to solution of an obstacle problem for a certain fully non-linear PDE. Consider

$$\sigma = \begin{bmatrix} 0, & 1 \\ 1, & 0 \end{bmatrix}.$$

Then we need $H_B \pm \sigma \geq 0$ in each point in D_Q . As we are looking for the best possible B satisfying these relationships, it is natural that we should require

$$\det(H_B + \sigma) = 0 \text{ or } \det(H_B - \sigma) = 0,$$

these are Monge–Ampère equations. **This approach is used in [59], [54]. Now we use another method to prove the existence of B_Q required in Theorem 2.7.**

2.1. More Bellman functions to prove the existence of Bellman function B_Q from Theorem 2.7. Dyadic shifts. We start with a much simpler “model” operator— T_σ . The logic will be the following. We want to get a sharp weighted estimate of $\|T_\sigma\|_{L^2(w) \rightarrow L^2(w)}$ via the A_2 characteristic of w . In the paper of Nazarov, Treil, Volberg, see [50], one can find that the norm $\|T_\sigma\|_{L^2(u) \rightarrow L^2(v)}$ is attained on some “simple” test functions—and that this holds for every pair u, v . Thus also for $u = v = w$. However, on the family \mathcal{T} of test functions one can compute the $N_{w,2}(T_\sigma) := \sup\{\|T_\sigma t\|_{L^2(w)} : t \in \mathcal{T}, \|t\|_{L^2(w)} = 1\}$. It turns out that

Theorem 2.9. $N_{w,2}(T_\sigma) \approx Q_{w,2}^{class}$.

J. Wittwer does that in [63] basing her approach on [50]; see also [56]. Thus, we get $\|T_\sigma\|_{L^2(w) \rightarrow L^2(w)} = N_{w,2}(T_\sigma) \approx Q_{w,2}^{class}$.

So let us show what the model operator is, what its sharp weighted estimate is, and how one obtains a special function (Bellman function) from this estimate.

Consider the family of dyadic singular operators T_σ :

$$T_\sigma f = \sum_{I \in \mathcal{D}} \sigma(I) (f, h_I) h_I.$$

Here \mathcal{D} is a dyadic lattice on \mathbb{R} , h_I is a Haar function associated with the dyadic interval I (h_I is normalized in $L^2(\mathbb{R}, dx)$), and $\sigma(I) = \pm 1$. We call the family T_σ *the martingale transform*. It is a dyadic analog of a Calderón–Zygmund operator. Here are important questions about T_σ , the first one about two-weight estimates and the second one about one weight estimates:

- 1) What are necessary and sufficient conditions for $\sup_\sigma \|T_\sigma\|_{L^2(u) \rightarrow L^2(v)} < \infty$?
- 2) What is the sharp bound on $\sup_\sigma \|T_\sigma\|_{L^2(w) \rightarrow L^2(w)}$ in terms of w ? How can one compute $\sup_\sigma \|T_\sigma\|_{L^2(w) \rightarrow L^2(w)}$?

These questions are dyadic analogs of notoriously difficult questions about “classical” Calderón–Zygmund operators like the Hilbert transform, the Riesz transforms and the Ahlfors–Beurling transform. The dyadic model is supposed to be easier than the continuous one. This turned out to be true. The answers to the questions above appeared in [50], [63]. Moreover these answers are key to answering questions about “classical” Calderón–Zygmund operators.

Strangely enough, the answer to the second question (which seems to be easier, because it is about “one weight”) seems to require the ideas from the “two-weight” case. Here is our explanation of this phenomena. The necessary and sufficient conditions on (u, v) to answer the first question were given in [50]. They amount to the fact that $\sup_\sigma \|T_\sigma\|_{L^2(u) \rightarrow L^2(v)}$ is almost attained on the family of simple test functions. This fact has beautiful consequences in the one weight case. For then $\sup_\sigma \|T_\sigma\|_{L^2(w) \rightarrow L^2(w)}$ is attainable (almost) on the family of simple test functions. One may try to compute $\sup_\sigma \|T_\sigma t\|_{L^2(w)}$ for every element of this test family, thus getting a good estimate for the norm $\sup_\sigma \|T_\sigma\|_{L^2(w) \rightarrow L^2(w)}$. Test functions are rather

simple, so this program can be carried out. This has been done in Wittwer's paper [63]. We will give another proof below. Here is the result. Recall that

$$Q_{w,2}^{dyadic} := \sup_{I \in \mathcal{D}} \langle w \rangle_I \langle w^{-1} \rangle_I.$$

Theorem 2.10.

$$\sup_{\sigma} \|T_{\sigma}\|_{L^2(w) \rightarrow L^2(w)} \leq A Q_{w,2}^{dyadic}.$$

Remark. We will postpone the proof of Theorem 2.10 (another use of **Bellman function technique**), here we will use it first to finish the proof of the existence of B_Q claimed in Theorem 2.6.

So we assume now that Theorem 2.10 is already proved. Let us rewrite Theorem 2.10 as follows

$$\sup_{\sigma(I)=\pm 1} |\Sigma_{I \in \mathcal{D}} \sigma(I) (f, h_I) (g, h_I)| \leq A Q_{w,2}^{dyadic} \|f\|_{L^2(w)} \|g\|_{L^2(w^{-1})},$$

or

$$\Sigma_{I \in \mathcal{D}} |(f, h_I)| |(g, h_I)| \leq A Q_{w,2}^{dyadic} \|f\|_{L^2(w)} \|g\|_{L^2(w^{-1})}.$$

This inequality is scaleless, so we write it as

$$(35) \quad J \in \mathcal{D}, \quad \frac{1}{4|J|} \Sigma_{I \in \mathcal{D}, I \subset J} |\langle f \rangle_{I_-} - \langle f \rangle_{I_+}| |\langle g \rangle_{I_-} - \langle g \rangle_{I_+}| |I| \leq A Q_{w,2}^{dyadic} \langle f^2 w \rangle_J^{1/2} \langle g^2 / w \rangle_J^{1/2}.$$

Here I_- , I_+ are the left and the right halves of I , and $\langle \cdot \rangle_l$ means averaging over l as usual. Given a fixed $J \in \mathcal{D}$ and a number $Q > 1$, we wish to introduce the Bellman function of (35):

$$\begin{aligned} B(X, Y, x, y, r, s) &= \sup \left\{ \frac{1}{4|J|} \Sigma_{I \in \mathcal{D}, I \subset J} |\langle f \rangle_{I_-} - \langle f \rangle_{I_+}| |\langle g \rangle_{I_-} - \langle g \rangle_{I_+}| |I| : \right. \\ &\quad \langle f \rangle_J = x, \langle g \rangle_J = y, \langle w \rangle_J = r, \langle w^{-1} \rangle_J = s, \\ &\quad \left. \langle f^2 w \rangle_J = X, \langle g^2 / w \rangle_J = Y, w \in A_2^{dyadic}, Q_{w,2}^{dyadic} \leq Q \right\}. \end{aligned}$$

Obviously, the function B does not depend on J , but it does depend on Q . Its domain of definition is the following:

$$R_Q := \{0 \leq (X, Y, x, y, r, s), x^2 \leq Xs, y^2 \leq Yr, 1 \leq rs \leq Q\}.$$

By (35) it satisfies

$$(36) \quad 0 \leq B \leq A Q X^{1/2} Y^{1/2}.$$

We are going to prove that it also satisfies the following “differential” inequality. Denote $v := (X, Y, x, y, r, s)$, $v_- = (X_-, Y_-, x_-, y_-, r_-, s_-)$, $v_+ = (X_+, Y_+, x_+, y_+, r_+, s_+)$, let v, v_+, v_- lie in R_Q , and let $v = \frac{1}{2}(v_- + v_+)$. Then

$$(37) \quad B(v) - \frac{1}{2} (B(v_+) + B(v_-)) \geq |x_+ - x_-| |y_+ - y_-|.$$

In fact, let f, g, w almost maximize $B(v)$ (on the interval J), let f_+, g_+, w_+ do this for $B(v_+)$, f_-, g_-, w_- do this for $B(v_-)$. The freedom of scale for B allows us to put f_+, g_+, w_+ on J_+ and f_-, g_-, w_- on J_- . Then we have “gargoyle” functions

$$F = \begin{cases} f_+ & \text{on } J_+ \\ f_- & \text{on } J_- \end{cases} \quad G = \begin{cases} g_+ & \text{on } J_+ \\ g_- & \text{on } J_- \end{cases} \quad W = \begin{cases} w_+ & \text{on } J_+ \\ w_- & \text{on } J_- \end{cases}.$$

Obviously, $\langle F \rangle_J = \frac{1}{2}(x_+ + x_-) = x$, $\langle G \rangle_J = y$, $\langle W \rangle_J = r$, $\langle W^{-1} \rangle_J = s$, $\langle F^2 W \rangle_J = X$, $\langle G^2 W^{-1} \rangle_J = Y$. These numbers together form the vector v . In other words F, G, W compete with f, g, w in the definition (35) of Bellman function $B(v)$. By this definition,

$$B(v) \geq \frac{1}{4|J|} \sum_{I \in \mathcal{D}, I \subset J} |\langle F \rangle_{I_-} - \langle F \rangle_{I_+}| |\langle G \rangle_{I_-} - \langle G \rangle_{I_+}| |I|.$$

But the almost optimality of f_+, g_+, w_+ on J_+ and f_-, g_-, w_- on J_- gives us (recall that $F = f_\pm$ on J_\pm , $G = g_\pm$ on J_\pm):

$$B(v_+) \leq \varepsilon + \frac{1}{4|J_+|} \sum_{I \in \mathcal{D}, I \subset J_+} |\langle F \rangle_{I_-} - \langle F \rangle_{I_+}| |\langle G \rangle_{I_-} - \langle G \rangle_{I_+}| |I|,$$

and

$$B(v_-) \leq \varepsilon + \frac{1}{4|J_-|} \sum_{I \in \mathcal{D}, I \subset J_-} |\langle F \rangle_{I_-} - \langle F \rangle_{I_+}| |\langle G \rangle_{I_-} - \langle G \rangle_{I_+}| |I|.$$

Combining these, we get

$$\begin{aligned} B(v) - \frac{1}{2}(B(v_+) + B(v_-)) &\geq -2\varepsilon + \frac{1}{4} |\langle F \rangle_{J_-} - \langle F \rangle_{J_+}| |\langle G \rangle_{J_-} - \langle G \rangle_{J_+}| \\ &= -2\varepsilon + \frac{1}{4} |\langle f_- \rangle_{J_-} - \langle f_+ \rangle_{J_+}| |\langle g_- \rangle_{J_-} - \langle g_+ \rangle_{J_+}| = -2\varepsilon + \frac{1}{4} |x_- - x_+| |y_- - y_+|. \end{aligned}$$

We are done with (37) because ε is an arbitrary positive number. Therefore, our B is a very concave function. We are going to modify B to have its Hessian satisfy the conclusion of Theorem 2.7. To do that we fix a compact K in the interior of R_Q , and we choose ε such that $100\varepsilon < \text{dist}(K, \partial R_Q)$. Consider the convolution of B with $\frac{1}{\varepsilon^6} \varphi(\frac{\cdot}{\varepsilon})$, $v \in \mathbb{R}^6$, where φ is a bell shape infinitely differentiable function with support in the unit ball of \mathbb{R}^6 . It is now very easy to see that this convolution (we call it $B_{K,Q}$) satisfies the following inequalities

$$(38) \quad 0 \leq B_{K,Q} \leq 6Q(X + Y),$$

and for any vector $\xi = (\xi_1, \xi_2, \xi_3, \xi_4, \xi_5, \xi_6) \in \mathbb{R}^6$,

$$(39) \quad -(d^2 B_{K,Q} \xi, \xi)_{\mathbb{R}^6} \geq 2|\xi_2| |\xi_3|.$$

The factor 2 appears because $B(v) - \frac{1}{2}(B(v_+) + B(v_-))$ in (37) corresponds to $-\frac{1}{2}d^2 B$, and $|x - x_+| = \frac{1}{2}|x_- - x_+|$ (the same being valid with y 's replacing x 's and $-$ replacing $+$).

Theorem 2.7 is completely proved modulo the proof of Theorem 2.10.

Proof of Theorem 2.10. To prove Theorem 2.10 we need the following decomposition:

Lemma 2.11.

$$(40) \quad h_I = \alpha_I h_I^w + \beta_I \frac{\chi_I}{\sqrt{I}},$$

where

- 1) $|\alpha_I| \leq \sqrt{\langle w \rangle_I}$,
- 2) $|\beta_I| \leq \frac{|\Delta_I w|}{\langle w \rangle_I}$, where $\Delta_I w := \langle w \rangle_{I_+} - \langle w \rangle_{I_-}$,
- 3) $\{h_I^w\}_I$ is an orthonormal basis in $L^2(w)$,
- 4) h_I^w assumes on I two constant values, one on I_+ and another on I_- .

Proof. To find α, β we first apply $\|\cdot\|_{L^2(w)}^2$ to both parts of (40): $\langle w \rangle_I = \|h_I\|_{L^2(w)}^2 = \alpha^2 + \beta^2 \langle w \rangle_I$, and secondly we multiply (40) by $\chi_I / \sqrt{|I|}$ and integrate with respect to $w dx$: $\frac{1}{2}(\langle w \rangle_{I_+} - \langle w \rangle_{I_-}) = \beta_I \langle w \rangle_I$. Clearly Lemma is proved. \square

Now let

$$\mathbb{S}F := \sum_I c_I(f, h_I) h_I, \text{ where constants } c_I \text{ are such that } |c_I| \leq 1.$$

Let $\sigma := w^{-1}$ for the rest of the proof. Fix $\phi \in L^2(w), \psi \in L^2(\sigma)$. We need to prove

$$(41) \quad |(\mathbb{S} \phi w, \psi \sigma)| \leq C \|\phi\|_w \|\psi\|_\sigma.$$

We estimate $(\mathbb{S} \phi w, \psi \sigma)$ as

$$\begin{aligned} & \left| \sum_I c_I(\phi w, h_I)(\psi \sigma, h_I) \right| \leq \\ & \sum_I |c_I(\phi w, h_I^w) \sqrt{\langle w \rangle_I} (\psi \sigma, h_I^\sigma) \sqrt{\langle \sigma \rangle_I}| + \\ & \sum_I |c_I \langle \phi w \rangle_I \frac{\Delta_I w}{\langle w \rangle_I} (\psi \sigma, h_I^\sigma) \sqrt{\langle \sigma \rangle_I} \sqrt{|I|}| + \\ & \sum_I |c_I \langle \psi \sigma \rangle_I \frac{\Delta_I \sigma}{\langle \sigma \rangle_I} (\phi w, h_I^w) \sqrt{\langle w \rangle_I} \sqrt{|I|}| + \\ & \sum_I |c_I \langle \phi w \rangle_I \langle \psi \sigma \rangle_I \frac{\Delta_I w}{\langle w \rangle_I} \frac{\Delta_I \sigma}{\langle \sigma \rangle_I} \sqrt{|I|} \sqrt{|I|}| =: I + II + III + IV. \end{aligned}$$

So we have

$$\begin{aligned} I & \leq \sum_I (\phi w, h_I^w) \sqrt{\langle w \rangle_I} \cdot (\psi \sigma, h_I^\sigma) \sqrt{\langle \sigma \rangle_I}, \quad II \leq \sum_I (\phi w, h_I^w) \sqrt{\langle w \rangle_I} \cdot \langle \psi \sigma \rangle_I \frac{|\Delta_I \sigma|}{\langle \sigma \rangle_I} \sqrt{|I|}, \\ III & \leq \sum_I \langle \phi w \rangle_I \frac{|\Delta_I w|}{\langle w \rangle_I} \sqrt{|I|} \cdot (\psi \sigma, h_I^\sigma) \sqrt{\langle \sigma \rangle_I}, \quad IV \leq \sum_I \langle \phi w \rangle_I \frac{|\Delta_I w|}{\langle w \rangle_I} \sqrt{|I|} \cdot \langle \psi \sigma \rangle_I \frac{|\Delta_I \sigma|}{\langle \sigma \rangle_I} \sqrt{|I|}. \end{aligned}$$

The estimate of I is trivial because h_I^w, h_I^σ are orthonormal systems in $L^2(w), L^2(\sigma)$ correspondingly:

$$(42) \quad I \leq \sup_I \sqrt{\langle w \rangle_I \langle \sigma \rangle_I} \sqrt{\sum_I (\phi w, h_I^w)^2} \sqrt{\sum_I (\psi \sigma, h_I^\sigma)^2} \leq [w]_{A_2}^{1/2} \|\phi\|_w \|\psi\|_\sigma.$$

To estimate the rest let us fix $\alpha \in (0, 1/2)$ and introduce

$$(43) \quad \mu_I := \langle w \rangle_I^\alpha \langle \sigma \rangle_I^\alpha \left(\frac{|\Delta_I w|^2}{\langle w \rangle_I^2} + \frac{|\Delta_I \sigma|^2}{\langle \sigma \rangle_I^2} \right) |I|.$$

We are going to give a **Bellman function** proof of the following lemma.

Lemma 2.12. *The sequence $\{\mu_I\}_{I \in D}$ is a Carleson sequence with Carleson constant at most $C[w]_{A_2}^\alpha$.*

We take Lemma 2.12 for granted till the end of the proof of Theorem 2.10. First introduce a notation, let μ be a positive measure on \mathbb{R} , then

$$M_\mu^d f(x) := \sup_{I \in D, x \in I} \frac{1}{\mu(I)} \int_I |f| d\mu.$$

This is called dyadic weighted maximal function. We will use it with $\mu = w dx$ or σdx .

To estimate IV , II , and symmetric to it III we notice that

$$\frac{|\Delta_I \sigma|}{\langle \sigma \rangle_I} \sqrt{|I|} \leq \langle w \rangle^{-\alpha/2} \langle \sigma \rangle^{-\alpha/2} \sqrt{\mu_I},$$

so, choosing $p \in (1, 2)$

$$\begin{aligned} \langle \psi \sigma \rangle_I \frac{|\Delta_I \sigma|}{\langle \sigma \rangle_I} \sqrt{|I|} &\leq \langle w \rangle^{-\alpha/2} \langle \sigma \rangle^{-\alpha/2} (\langle |\psi|^p \sigma \rangle_I)^{1/p} \langle \sigma \rangle^{1-1/p} \sqrt{\mu_I} \leq \\ &\langle w \rangle^{-\alpha/2} \langle \sigma \rangle_I^{1-\alpha/2} \inf_{x \in I} (M_\sigma^d |\psi|^p(x))^{1/p} \cdot \sqrt{\mu_I}, \end{aligned}$$

where M_σ^d is the dyadic weighted maximal function. Therefore,

$$\begin{aligned} IV &\leq \sum_I \langle w \rangle^{1-\alpha} \langle \sigma \rangle_I^{1-\alpha} \inf_I (M_\sigma^d |\psi|^p)^{1/p} \cdot \inf_I (M_w^d |\phi|^p)^{1/p} \cdot \mu_I. \\ II &\leq \sum_I (\phi w, h_I^w) \langle w \rangle_I^{1-\alpha/2} \langle \sigma \rangle_I^{1-\alpha/2} \frac{\inf_I (M_\sigma^d |\psi|^p)^{1/p}}{\langle w \rangle_I^{1/2}} \cdot \sqrt{\mu_I}. \end{aligned}$$

The estimate of III will be totally symmetric, so we omit it. We continue:

$$\begin{aligned} IV &\leq [w]_{A_2}^{1-\alpha} \sum_I \inf_I (M_\sigma^d |\psi|^p)^{1/p} \cdot \inf_I (M_w^d |\phi|^p)^{1/p} \cdot \mu_I. \\ II &\leq [w]_{A_2}^{1-\alpha/2} \sqrt{\sum_I (\phi w, h_I^w)^2} \sqrt{\sum_I \frac{\inf_I (M_\sigma^d |\psi|^p)^{2/p}}{\langle w \rangle_I}} \cdot \mu_I. \end{aligned}$$

Choose $F = (M_\sigma^d |\psi|^p)^{1/p} \cdot (M_w^d |\phi|^p)^{1/p}$ and $G = (M_\sigma^d |\psi|^p)^{2/p}$ and apply the following simple lemma (**Exercise!**)

Lemma 2.13. *Let $\{\alpha_L\}_{L \in D}$ define Carleson measure with intensity B . Let F be a positive function on the line. Then*

$$(44) \quad \sum_L (\inf_L F) \alpha_L \leq 2B \int_{\mathbb{R}} F dx.$$

$$(45) \quad \sum_L \frac{\inf_L G}{\langle w \rangle_L} \alpha_L \leq C B \int_{\mathbb{R}} \frac{G}{w} dx.$$

Then using Lemma 2.12 we get

$$\begin{aligned} IV &\leq [w]_{A_2}^{1-\alpha} [w]_{A_2}^{\alpha} \int_{\mathbb{R}} (M_{\sigma}^d |\psi|^p)^{1/p} \cdot (M_w^d |\phi|^p)^{1/p} dx = \\ &[w]_{A_2} \int_{\mathbb{R}} (M_{\sigma}^d |\psi|^p)^{1/p} \cdot (M_w^d |\phi|^p)^{1/p} w^{1/2} \sigma^{1/2} dx \leq \\ &[w]_{A_2} \sqrt{\int_{\mathbb{R}} (M_w^d |\phi|^p)^{2/p} w dx} \sqrt{\int_{\mathbb{R}} (M_{\sigma}^d |\psi|^p)^{2/p} \sigma dx} \leq C [w]_{A_2} \|\phi\|_w \|\psi\|_{\sigma}. \end{aligned}$$

As to II , we have again using Lemma 2.13 (the second part) and Lemma 2.12:

$$\begin{aligned} II &\leq C [w]_{A_2}^{1-\alpha/2} [w]_{A_2}^{\alpha/2} \sqrt{\sum_I (\phi w, h_I^w)^2} \sqrt{\int_{\mathbb{R}} \frac{(M_{\sigma}^d |\psi|^p)^{2/p}(x)}{w(x)} dx} \leq \\ &C [w]_{A_2} \|\phi\|_w \sqrt{\int_{\mathbb{R}} (M_{\sigma}^d |\psi|^p)^{2/p}(x) \sigma(x) dx} \leq C [w]_{A_2} \|\phi\|_w \|\psi\|_{\sigma}. \end{aligned}$$

Theorem 2.10 is completely proved, function B_Q is constructed. □

We need only to see the validity of Lemma 2.12. This is done by yet another **Bellman function**.

Bellman proof of Lemma 2.12. We prove even a more general statement, namely we prove the version in \mathbb{R}^d and even in each metric space with **geometric doubling condition** and doubling measure μ . So let us have a metric space with geometric doubling condition, meaning that every ball of radius r can fit only at most K disjoint balls of radius $r/2$, K being independent of the ball and its radius. Such metric spaces carry a doubling measure μ by a theorem of Konyagin–Volberg [48], and let D denote the family of “dyadic cubes” on this metric space (constructions are numerous, the first belongs to M. Christ [26]), and let $s_i(I)$ are dyadic children of $I \in D$. Finally, let $I \in D$ and let

$$\mu_I := (\langle w \rangle_{\mu, I} \langle \sigma \rangle_{\mu, I})^{\alpha} \left(\frac{(\langle w \rangle_{\mu, s_i(I)} - \langle w \rangle_{\mu, I})^2}{\langle w \rangle_{\mu, I}^2} + \frac{(\langle \sigma \rangle_{\mu, s_i(I)} - \langle \sigma \rangle_{\mu, I})^2}{\langle \sigma \rangle_{\mu, I}^2} \right) \mu(I).$$

Lemma 2.12 becomes the following statement, which we are proving below:

$$(46) \quad \forall I \in D \quad \sum_{J \in D, J \subset I} \mu_J \leq C_{\alpha} [w]_{\mu, A_2}^{\alpha} \mu(I).$$

Let $Q > 1, 0 < \alpha < \frac{1}{2}$. In domain $\Omega_Q := \{(x, y) : x > 0, y > 0, 1 < xy \leq Q\}$ function $b_Q(x, y) := x^\alpha y^\alpha$ satisfies the following estimate of its Hessian matrix (of its second differential form, actually)

$$-d^2 b_Q(x, y) \geq \alpha(1 - 2\alpha)x^\alpha y^\alpha \left(\frac{(dx)^2}{x^2} + \frac{(dy)^2}{y^2} \right).$$

The form $-d^2 b_Q(x, y) \geq 0$ everywhere in $x > 0, y > 0$. Also obviously $0 \leq b_Q(x, y) \leq Q^\alpha$ in Ω_Q .

Proof. Direct calculation. □

Fix now a cube I and let $s_i(I), i = 1, \dots, M$, be all its sons. Let $a = (\langle w \rangle_{\mu, I}, \langle \sigma \rangle_{\mu, I})$, $b_i = (\langle w \rangle_{\mu, s_i(I)}, \langle \sigma \rangle_{\mu, s_i(I)})$, $i = 1, \dots, M$, be points—obviously—in Ω_Q , where Q temporarily means $[w]_{A_2}$. Consider $c_i(t) = a(1-t) + b_i t$, $0 \leq t \leq 1$ and $q_i(t) := b_Q(c_i(t))$. We want to use Taylor's formula

$$(47) \quad q_i(0) - q_i(1) = -q'_i(0) - \int_0^1 dx \int_0^x q''_i(t) dt.$$

Notice two things: Sublemma shows that $-q''_i(t) \geq 0$ always. Moreover, it shows that if $t \in [0, 1/2]$, then we have that the following qualitative estimate holds:

$$(48) \quad -q''_i(t) \geq c(\langle w \rangle_{\mu, I} \langle \sigma \rangle_{\mu, I})^\alpha \left(\frac{(\langle w \rangle_{\mu, s_i(I)} - \langle w \rangle_{\mu, I})^2}{\langle w \rangle_{\mu, I}^2} + \frac{(\langle \sigma \rangle_{\mu, s_i(I)} - \langle \sigma \rangle_{\mu, I})^2}{\langle \sigma \rangle_{\mu, I}^2} \right)$$

This requires a small explanation. If we are on the segment $[a, b_i]$, then the first coordinate of such a point cannot be larger than $C \langle w \rangle_{\mu, I}$, where C depends only on doubling of μ (not w). This is obvious. The same is true for the second coordinate with the obvious change of w to σ . But there is no such type of estimate from below on this segment: the first coordinate cannot be smaller than $k \langle w \rangle_{\mu, I}$, but k may (and will) depend on the doubling of w (so ultimately on its $[w]_{A_2}$ norm). In fact, at the “right” endpoint of $[a, b_i]$ the first coordinate is $\langle w \rangle_{\mu, s_i(I)} \leq \int_I w d\mu / \mu(s_i(I)) \leq C \int_I w d\mu / \mu(I) = C \langle w \rangle_{\mu, I}$, with C only depending on the doubling of μ . But the estimate from below will involve the doubling of w , which we must avoid. But if $t \in [0, 1/2]$, and we are on the “left half” of interval $[a, b_i]$ then obviously the first coordinate is $\geq \frac{1}{2} \langle w \rangle_{\mu, I}$ and the second coordinate is $\geq \frac{1}{2} \langle \sigma \rangle_{\mu, I}$.

We do not need to integrate $-q''_i(t)$ for all $t \in [0, 1]$ in (47). We can only use integration over $[0, 1/2]$ noticing that $-q''_i(t) \geq 0$ otherwise. Then the chain rule

$$q''_i(t) = (b_Q(c_i(t)))'' = (d^2 b_Q(c_i(t)))(b_i - a, b_i - a),$$

(where (\cdot, \cdot) means the usual scalar product in \mathbb{R}^2) immediately gives us (48) with constant c depending on the doubling of μ but *independent* of the doubling of w .

Next step is to add all (47), with convex coefficients $\frac{\mu(s_i(I))}{\mu(I)}$, and to notice that $\sum_{i=1}^M \frac{\mu(s_i(I))}{\mu(I)} q'_i(0) = \nabla b_Q(a) \sum_{i=1}^M (a - b_i) \frac{\mu(s_i(I))}{\mu(I)} = 0$, because by definition

$$a = \sum_{i=1}^M \frac{\mu(s_i(I))}{\mu(I)} b_i.$$

Notice that the addition of all (47), with convex coefficients $\frac{\mu(s_i(I))}{\mu(I)}$ gives us now (we take into account (48) and positivity of $-q_i''(t)$)

$$b_Q(a) - \sum_{i=1}^M \frac{\mu(s_i(I))}{\mu(I)} b_Q(b_i) \geq c c_1 (\langle w \rangle_{\mu,I} \langle \sigma \rangle_{\mu,I})^\alpha \sum_{i=1}^M \left(\frac{(\langle w \rangle_{\mu,s_i(I)} - \langle w \rangle_{\mu,I})^2}{\langle w \rangle_{\mu,I}^2} + \frac{(\langle \sigma \rangle_{\mu,s_i(I)} - \langle \sigma \rangle_{\mu,I})^2}{\langle \sigma \rangle_{\mu,I}^2} \right).$$

We used here the doubling of μ again, by noticing that $\frac{\mu(s_i(I))}{\mu(I)} \geq c_1$ (recall that $s_i(I)$ and I are almost balls of comparable radii). We rewrite the previous inequality using our definition of $\Delta_I w, \Delta_I \sigma$ listed above as follows

$$\mu(I) b_Q(a) - \sum_{i=1}^M \mu(s_i(I)) b_Q(b_i) \geq c c_1 (\langle w \rangle_{\mu,I} \langle \sigma \rangle_{\mu,I})^\alpha \left(\frac{(\Delta_I w)^2}{\langle w \rangle_{\mu,I}^2} + \frac{(\Delta_I \sigma)^2}{\langle \sigma \rangle_{\mu,I}^2} \right) \mu(I).$$

Notice that $b_Q(a) = \langle w \rangle_{\mu,I}^\alpha \langle \sigma \rangle_{\mu,I}^\alpha$. Now we iterate the above inequality and get for any of dyadic I 's:

$$\sum_{J \subset I, J \in D} (\langle w \rangle_{\mu,J} \langle \sigma \rangle_{\mu,J})^\alpha \left(\frac{(\Delta_J w)^2}{\langle w \rangle_{\mu,J}^2} + \frac{(\Delta_J \sigma)^2}{\langle \sigma \rangle_{\mu,J}^2} \right) \mu(J) \leq C Q^\alpha \mu(I).$$

This is exactly the Carleson property of the measure $\{\mu_I\}$ indicated in our Lemma 2.12, with Carleson constant $C Q^\alpha$. The proof showed that C depended only on $\alpha \in (0, 1/2)$ and on the doubling constant of measure μ . Lemma 2.12 is completely proved. □

3. ESTIMATES FOR AHLFORS–BEURLING OPERATOR. TOWARDS THE BIG IWANIEC PROBLEM BY BELLMAN FOOTSTEPS

In the previous section we estimated AB operator T in weighted $L^2(w)$. The estimate was sharp in $[w]_{A_2}$:

$$(49) \quad |(Tf, g)| \leq C [w]_{A_2} \|f\|_{L^2(w)} \|g\|_{L^2(w^{-1})},$$

it implied a sharp in $[w]_{A_p}$ estimate in weighted $L^p(w)$:

$$(50) \quad |(Tf, g)| \leq C(p) [w]_{A_p}^{\max(1, \frac{1}{p-1})} \|f\|_{L^p(w)} \|g\|_{L^{p'}(w^{-1/(p-1)})}, \quad p' := p/(p-1).$$

But we did not care about $C, C(p)$ at all. Now we consider just $w = 1$, but we care about $C(p)$ very much. Big Iwaniec's problem conjectures

$$(51) \quad C(p) = \max(p, p/(p-1)) - 1 =: p^* - 1.$$

This is **open** at the moment of writing this phrase. Using various **Bellman functions** we will show the row of improvements

$$(52) \quad C(p) \leq 2(p^* - 1).$$

$$(53) \quad C(p) \leq 1.7(p^* - 1).$$

$$(54) \quad C(p) \leq 1.575(p^* - 1).$$

$$(55) \quad C(p) \leq 1.4(p^* - 1).$$

Recall things that we already know: 1) $T = R_1^2 - R_2^2 + 2iR_1R_2$, where R_i are Riesz transforms = multipliers with symbol $\xi_i/(|\xi_1|^2 + |\xi_2|^2)^{1/2}$, $i = 1, 2$;

$$(56) \quad 2) (R_i^2 f, \bar{g}) = -2 \iint_{\mathbb{R}_+^3} \frac{\partial f}{\partial x_i} \frac{\partial g}{\partial x_i} dx dt,$$

where f, g in the left hand side are from $C_0^\infty(\mathbb{R}^2)$, and f, g in the right hand side are heat extensions of functions in the left.

Hence Conjecture 51 is nothing else as the following innocent looking **conjecture**

$$(57) \quad 2 \left| \iint_{\mathbb{R}_+^3} \left(\frac{\partial f}{\partial x_1} + i \frac{\partial f}{\partial x_2} \right) \cdot \left(\frac{\partial g}{\partial x_1} + i \frac{\partial g}{\partial x_2} \right) dx dt \right| \leq (p-1) \|f\|_p \|g\|_{p'}, \quad p > 2.$$

Let complex-valued functions $f = u + iv, g = \phi + i\psi$. Consider $f := (u, v)$ as a map $\mathbb{R}^2 \rightarrow \mathbb{R}^2$, do the same with $G := (\phi, \psi)$. We have Jacobian matrices DF, DG then. These are 2×2 matrices.

Imagine that we want to have a stronger estimate than (57) (which is probably **too much!**):

$$(58) \quad 2 \iint_{\mathbb{R}_+^3} \left| \frac{\partial f}{\partial x_1} + i \frac{\partial f}{\partial x_2} \right| \cdot \left| \frac{\partial g}{\partial x_1} + i \frac{\partial g}{\partial x_2} \right| dx dt \leq (p-1) \|f\|_p \|g\|_{p'}, \quad p > 2.$$

This is exactly

$$(59) \quad 2 \iint_{\mathbb{R}_+^3} (|DF|_2^2 - 2 \det DF)^{1/2} (|DG|_2^2 - 2 \det DG)^{1/2} dx dt \leq (p-1) \|f\|_p \|g\|_{p'}, \quad p > 2,$$

where $|\cdot|_2$ is the Hilbert-Schmidt norm of the matrix.

Nobody can prove (59) and equivalent to it (58). They may be wrong!

However, we will start with proving slightly lighter estimates:

$$(60) \quad 2 \iint_{\mathbb{R}_+^3} \left| \frac{\partial f}{\partial x_1} \right| \left| \frac{\partial g}{\partial x_1} \right| + \left| \frac{\partial f}{\partial x_2} \right| \left| \frac{\partial g}{\partial x_2} \right| dx dt \leq (p-1) \|f\|_p \|g\|_{p'}, \quad p > 2.$$

Moreover, we will prove a stronger than (60) (but weaker than (58)) estimate

$$(61) \quad 2 \iint_{\mathbb{R}_+^3} \left(\left| \frac{\partial f}{\partial x_1} \right|^2 + \left| \frac{\partial f}{\partial x_2} \right|^2 \right)^{1/2} \left(\left| \frac{\partial g}{\partial x_1} \right|^2 + \left| \frac{\partial g}{\partial x_2} \right|^2 \right)^{1/2} dx dt \leq (p-1) \|f\|_p \|g\|_{p'}, \quad p > 2.$$

This will give us (52), (53) correspondingly. To get to (54) and further improvements as (55) we will need a bit more (stochastic integrals).

Notice that we already know (by (56)) that (60) immediately proves the following

Theorem 3.1. $\|R_1^2 - R_2^2\|_p \leq p-1, \quad p \geq 2.$

Because $2R_1R_2 = U \circ (R_1^2 - R_2^2) \circ U^{-1}$, where U is an isometry in all L^p spaces (in fact, U is the rotation of the argument of function by 45°), we get (52) from doubling the claim of Theorem 3.1.

proof of (60). The first step is by examination of what we already had in Section 2.1 after the statement of Theorem 2.10. We do now **exactly** the same:

Suppose we have the following inequality for functions on interval $[0, 1]$ provided with dyadic lattice \mathcal{D} :

$$(62) \quad \Sigma_{I \in \mathcal{D}} |(f, h_I)| |(g, h_I)| \leq (p-1) \|f\|_{L^p} \|g\|_{L^{p'}}, p \geq 2.$$

This inequality is scaleless, so we write it as

$$(63) \quad J \in \mathcal{D}, \quad \frac{1}{4|J|} \Sigma_{I \in \mathcal{D}, I \subset J} |\langle f \rangle_{I_-} - \langle f \rangle_{I_+}| |\langle g \rangle_{I_-} - \langle g \rangle_{I_+}| |I| \leq (p-1) \langle |f|^p \rangle_J^{1/p} \langle |g|^{p'} \rangle_J^{1/p'}.$$

Here I_-, I_+ are the left and the right halves of I , and $\langle \cdot \rangle_l$ means averaging over l as usual. Given a fixed $J \in \mathcal{D}$, $p \geq 2$, we wish to introduce the Bellman function of (35):

$$B_p(X, Y, x, y) = \sup \left\{ \frac{1}{4|J|} \Sigma_{I \in \mathcal{D}, I \subset J} |\langle f \rangle_{I_-} - \langle f \rangle_{I_+}| |\langle g \rangle_{I_-} - \langle g \rangle_{I_+}| |I| : \right. \\ \left. \langle f \rangle_J = x, \langle g \rangle_J = y, \langle |f|^p \rangle_J = X, \langle |g|^{p'} \rangle_J = Y \right\}.$$

Obviously, the function B does not depend on J , but it does depend on p . Its domain of definition is the following:

$$R_p := \{(X, Y, \mathbf{x}, \mathbf{y}), |\mathbf{x}|^p \leq X, |\mathbf{y}|^{p'} \leq Y\}.$$

By (62) it satisfies

$$(64) \quad 0 \leq B \leq (p-1) X^{1/p} Y^{1/p'}.$$

We are going to prove that it also satisfies the following “differential” inequality. Denote $v := (X, Y, \mathbf{x}, \mathbf{y})$, $v_- = (X_-, Y_-, \mathbf{x}_-, \mathbf{y}_-)$, $v_+ = (X_+, Y_+, \mathbf{x}_+, \mathbf{y}_+)$, let v, v_+, v_- lie in R_p , and let $v = \frac{1}{2}(v_- + v_+)$. Then

$$(65) \quad B(v) - \frac{1}{2} (B(v_+) + B(v_-)) \geq \frac{1}{4} |\mathbf{x}_+ - \mathbf{x}_-| |\mathbf{y}_+ - \mathbf{y}_-|.$$

The proof is **verbatim** the same as in Section 2.1. And this inequality in infinitesimal sense becomes

$$(66) \quad d^2 B_p \geq 2 |d\mathbf{x}| |d\mathbf{y}|.$$

Having the function B_p satisfying

- 1) $0 \leq B_p \leq (p-1) X^{1/p} Y^{1/p'}$;
- 2) $-d^2 B_p \geq 2 |d\mathbf{x}| |d\mathbf{y}|$.

Assuming that B_p is sufficiently smooth (which incidentally it is, one can write the formula for B_p), we can repeat **verbatim** we can repeat the proof of Theorem 2.6: we start with analyzing ($x = (x_1, x_2) \in \mathbb{R}^2$)

$$(67) \quad \left(\frac{\partial}{\partial t} - \Delta \right) b(x, t)$$

exactly as in the proof of Theorem 2.6: the only difference that b now is not $B_Q \circ v$ but our $B_p \circ v$ and v also slightly different, it is now

$$v(x, t) := (|f|^p(x, t), |g|^{p'}(x, t), f(x, t), g(x, t)),$$

where these are heat extensions of functions on \mathbb{R}^2 with corresponding symbol. We estimate the expression in (67) in a pointwise way from below using 2), and in the average on a slab, using 1) we got exactly (60), Theorem 3.1, and, therefore, (52).

Remark. Notice that variables \mathbf{x}, \mathbf{y} are complex, they are “bench guards” (“mestoblyustiteli”) for complex-valued functions $f = u + iv, g = \phi + i\psi$. So actually B_p is a function of 6 real variables, and, hence, (66) should be understood as

$$(68) \quad -d^2 B_p(X; Y; u, v; \phi, \psi) = (H_{B_p} h, h) \geq 2\sqrt{du^2 + dv^2} \sqrt{d\phi^2 + d\psi^2},$$

where u, v, ϕ, ψ are just real variables (they are “bench guards” for functions with the same symbols and their heat extensions), and $h = (dX, dY, du, dv, d\phi, d\psi)$ is a notation (strange may be) for an arbitrary vector in \mathbb{R}^6 . \square

To obtain (53) we notice first that in Theorem 3.1 we can use $R_1 \cos \theta - R_2 \sin \theta$ in place of R_1 , and $R_1 \sin \theta + R_2 \cos \theta$ in place of R_2 . In fact this is just application of rotation on θ in arguments. Then we notice that $(R_1 \cos \theta - R_2 \sin \theta)^2 - (R_1 \sin \theta + R_2 \cos \theta)^2 = (R_1^2 - R_2^2) \cos 2\theta - 2R_1 R_2 \sin 2\theta$. Therefore, we got

Theorem 3.2. *For any $\phi \in (0, 2\pi]$, $\|(R_1^2 - R_2^2) \cos \phi - 2R_1 R_2 \sin \phi\|_p \leq p - 1$ if $p \geq 2$.*

We notice that a certain estimate of $T = (R_1^2 - R_2^2) + 2iR_1 R_2$ can be obtained if we answer the following question. Suppose A, B are two operators in $L^p(\mu)$, and for any angle $\|A \cos \phi - B \sin \phi\|_p \leq 1$, then what is the estimate of $\|A - iB\|_p$?

This is easy on real functions, let $f \in L^p_{real}(\mu)$, and let A, B map real functions to real functions ($A = R_1^2 - R_2^2, B = 2R_1 R_2$ are such). In fact,

$$\begin{aligned} \int |f|^p d\mu &\geq \int |(Af)(x) \cos \phi + (Bf)(x) \sin \phi|^p d\mu = \\ &\int (|Af|^2 + |Bf|^2)^{p/2} |\cos(a(x) - \phi)|^p d\mu(x). \end{aligned}$$

Integrate this over $\frac{1}{2\pi} \int_0^{2\pi} \dots$, by Fubini' theorem we will get

$$(69) \quad \int |f|^p d\mu \geq \int (|Af|^2 + |Bf|^2)^{p/2} d\mu \cdot \frac{1}{2\pi} \int_0^{2\pi} |\cos \phi|^p d\phi.$$

Put

$$\tau(p) := \left(\frac{1}{2\pi} \int_0^{2\pi} |\cos \phi|^p d\phi \right)^{1/p},$$

then on **real functions**

$$(70) \quad \|A + iB\|_p \leq \sup_{\phi} \|A \cos \phi + B \sin \phi\|_p / \tau(p).$$

Unfortunately this was in real category. We do not know how obtain (70)—or something like that—for general operators A, B on complex function. May be this is also **an exercise?**

However, we will obtain now (53). First we need

The proof of (61) . We use the following elementary lemma from Linear Algebra:

Lemma 3.3 (Linear Algebra lemma). *Let A, B, C be nonnegative matrices of size $d \times d$. Let*

$$(71) \quad (Ah, h) \geq 2(Bh, h)^{1/2}(Ch, h)^{1/2}, \quad \forall h \in \mathbb{C}^d.$$

Then there exists $\tau \in (0, \infty)$ independent of h such that

$$(Ah, h) \geq \tau(Bh, h) + \frac{1}{\tau}(Ch, h), \quad \forall h \in \mathbb{C}^d.$$

Proof. Exercise. □

We apply this lemma separately to h_1, h_2 , where $(f = u + iv, g = \phi + i\psi,)$

$$h_1 = (\partial_{x_1}|f|^p(x, t), \partial_{x_1}|g|^{p'}(x, t), \partial_{x_1}u(x, t), \partial_{x_1}v(x, t), \partial_{x_1}\phi(x, t), \partial_{x_1}\psi(x, t)),$$

$$h_2 = (\partial_{x_2}|f|^p(x, t), \partial_{x_2}|g|^{p'}(x, t), \partial_{x_2}u(x, t), \partial_{x_2}v(x, t), \partial_{x_2}\phi(x, t), \partial_{x_2}\psi(x, t)),$$

and $A = H_{B_p}(|f|^p(x, t), |g|^{p'}(x, t), u(x, t), v(x, t), \phi(x, t), \psi(x, t))$, and B consisting of all zeros except 3, 3 and 4, 4 entries, where we have 1, and C consisting of all zeros except 5, 5 and 6, 6 entries, where we have 1.

Then we immediately get (61). □

The proof of (53) . We use the previous notations. We want a better estimate of $Tf = (A + iB)(u + iv) = Au - Bv + i(Av + Bu)$. Using the trick above (69) we can average the following equality over $(0, 2\pi)$

$$\begin{aligned} & \int |(Au - Bv)(x) \cos \phi + (Av + Bu)(x) \sin \phi|^p d\mu = \\ & \int (|Au - Bv|^2 + |Av + Bu|^2)^{\frac{p}{2}} |\cos(a(x) - \phi)|^p d\mu(x). \end{aligned}$$

Then we get

$$\begin{aligned} \tau(p) \cdot \left(\int |Tf|^p \right)^{1/p} & \leq \sup_{\phi} \left(\int |(Au - Bv)(x) \cos \phi + (Av + Bu)(x) \sin \phi|^p \right)^{1/p} = \\ & \sup_{\phi} \sup_{\text{real } \psi, \|\psi\|_{p'} \leq 1} \int [(Au - Bv)(x) \cos \phi + (Av + Bu)(x) \sin \phi] \psi(x) dx =: E \end{aligned}$$

However the last expression can be rewritten using (56) and integration by parts as follows:

$$\begin{aligned} E & = 2\Re \iint_{\mathbb{R}_+^3} (\partial_{x_1} + i\partial_{x_2})f(x, t)(\partial_{x_1} + i\partial_{x_2})e^{-i\phi}\psi(x, t)dxdt \leq \\ & 2\sqrt{2} \iint (|\partial_{x_1}f|^2 + |\partial_{x_2}f|^2)^{1/2} ((\partial_{x_1}\psi)^2 + (\partial_{x_2}\psi)^2)^{1/2} \leq \sqrt{2}(p-1)\|f\|_p, \quad p > 2. \end{aligned}$$

We used (61). Here $\sqrt{2}$ appeared trivially from

$$|(\partial_{x_1} + i\partial_{x_2})f(x, t)| \leq \sqrt{2} (|\partial_{x_1}f|^2 + |\partial_{x_2}f|^2)^{1/2}.$$

Finally we get

$$(72) \quad \|T\|_p \leq \frac{\sqrt{2}(p-1)}{\left(\frac{1}{2\pi} \int_0^{2\pi} |\cos \phi|^p d\phi \right)^{1/p}}, \quad p > 2.$$

Asymptotically this is $1.41\dots(p-1)$. Choosing large p , interpolating between L^2 , where the norm of T is 1 and the estimate (72) for this large p , then optimizing by the choice of p one can get (53) (**exercise!**). \square

Notice that (61) immediately proves the following

Theorem 3.4. 1) $\|T : L_{real}^p \rightarrow L^p\| \leq \sqrt{2}(p-1)$, $p \geq 2$;
 2) $|(Tf, g)| \leq (p-1)\|f\|_p\|g\|_{p'}$, $p \geq 2$, if f, g are real valued.

Proof. Just look at (59), compare it with (61) and the fact that $(|DF|_2^2 - 2 \det DF)^{1/2} \leq \sqrt{2}|DF|$, $F = (u, v)$. We also need to notice that in this inequality for real valued $f = u + i0$ we have $\det DF = 0$ and the constant $\sqrt{2}$ can be replaced by 1. Finish the proof: **exercise**. \square

So everything above hinges on inequality (62). This inequality was proved by Burkholder in mid 80's and it is one of the **remarkable inventions**. It is done by use of **Bellman function technique**.

3.1. The proof of inequality (62). Burkholder's Bellman function. We follow [19], [23], [25]—but loosely. See also the exposition in the review paper [4].

Let f be real valued on $[0, 1] =: I_0$. Let $\{h_I\}_{I \in \mathcal{D}}$ be the usual Haar functions on I_0 normalized in L^2 . Consider an operator

$$T_\varepsilon f = \sum_{I \in \mathcal{D}} \varepsilon_I(f, h_I) h_I, \quad \varepsilon := \{\varepsilon_I\}_I, \quad \varepsilon_I = \pm 1.$$

This family is called martingale transforms.

Burkholder proved the following remarkable

Theorem 3.5. $\sup_\varepsilon \|T_\varepsilon\|_p = p^* - 1 := \max(p, p/(p-1)) - 1$.

He gave several proofs, all difficult, to be found in [19]–[25]. Another proof by Vasyunin–Volberg see arxiv: 1006.2633, [59].

In all these proofs the following object is indispensable. It is Burkholder's **Bellman function**.

Let $\Omega := \{(x, y, z) : |x|^p \leq z\}$ and let

$$B(x, y, z) := \sup \{ \|g\|_p^p : \langle f \rangle_{I_0} = x, \langle g \rangle_{I_0} = y, \langle |f|^p \rangle_{I_0} = z, \forall I \in \mathcal{D} \mid (g, h_I) = (f, h_I) \}.$$

Symmetries:

$$(73) \quad B(tx, ty, t^p z) = t^p B(x, y, z), \quad B(-x, y) = B(x, y), \quad B(x, -y) = B(x, y).$$

Burkholder found the formula for B :

Consider for positive x, y

$$F_p(x, y) = \begin{cases} y^p - (p^* - 1)^p x^p, & \text{if } y \leq (p^* - 1)x; \\ p \left(1 - \frac{1}{p^*}\right)^{p-1} (y+x)^{p-1} (y - (p^* - 1)x), & \text{if } y \geq (p^* - 1)x. \end{cases}$$

Consider the solution of an implicit equation:

$$F_p(|x|, |y|) = F_p(z^{1/p}, B^{1/p}(x, y, z)).$$

If $p \geq 2$ Burkholder's function is the solution of this equation. If $p \in (1, 2]$, then one considers $F_p(|y|, |x|) = F_p(B^{1/p}(x, y, z), z^{1/p})$.

Obviously one gets a

Theorem 3.6. $B(0, 0, 1) = (p^* - 1)^p$,

which gives Theorem 3.5, from which we get that (62) is proved right away. In fact,

The proof of (62). We write $\sup_{\varepsilon} |(T_{\varepsilon} f, g)| \leq (p^* - 1) \|f\|_p \|g\|_{p'}$, which follows from Theorem 3.5. But this supremum obviously is equal to

$$\Sigma_{I \in \mathcal{D}} |(f, h_I)| |(g, h_I)|.$$

Therefore (62) is proved. □

Remarks. 1) As soon as (62) is proved we have our **Bellman** function B_p .

2) It gives all our inequalities like (61) and its consequences like (53).

3) It is **not** Burkholder's function.

4) The existence of our Bellman function B_p follows from the existence of Burkholder's Bellman function. **These are demographic creatures, they create one another—we saw this in previous sections too.**

We are left to prove Theorem 3.5. Instead of finding exact formula for $B(x, y, z)$ listed above we will use a certain shortcut (invented already by Burkholder himself). Suppose Burkholder's B is finite.

The shortcut proof of Theorem 3.5. Along with symmetries (73) it has very good concavity properties:

$$(74) \quad B(x, y, z) - \frac{1}{2}(B(x + \alpha, y + \alpha, z + \beta) + B(x - \alpha, y - \alpha, z - \beta)) \geq 0,$$

if all points lie in Ω . Also

$$(75) \quad B(x, y, z) - \frac{1}{2}(B(x + \alpha, y - \alpha, z + \beta) + B(x - \alpha, y + \alpha, z - \beta)) \geq 0,$$

if all points lie in Ω .

Inequalities (74), (75) are left as **exercise**.

Notice that this means that

$$M(a, b, c) := B(a + b, a - b, c)$$

is concave in (a, c) , and in (b, c) .

Definition. Such M is called bi-concave.

Definition. Function φ on \mathbb{R}^2 is called zigzag concave if

$$\varphi(x, y) - \frac{1}{2}(\varphi(x + \alpha, y + \alpha) + \varphi(x - \alpha, y - \alpha)) \geq 0,$$

$$\varphi(x, y) - \frac{1}{2}(\varphi(x + \alpha, y - \alpha) + \varphi(x - \alpha, y + \alpha)) \geq 0,$$

or, which is the same as,

$$\varphi(x, y) - \frac{1}{2}(\varphi(x^+, y^+) + \varphi(x^-, y^-)) \geq 0, \text{ if}$$

$$|x^+ - x^-| = |y^+ - y^-|, x = \frac{1}{2}(x^+ + x^-), y = \frac{1}{2}(y^+ + y^-).$$

Theorem 3.7. Put $\varphi(x, y) := \sup_{(x, y, z) \in \Omega} [B(x, y, z) - (p^* - 1)^p z]$. It is zigzag concave. It is the least zigzag concave majorant of $h(x, y) := |y|^p - (p^* - 1)^p |x|^p$. There is **no** zigzag concave majorant ψ such that $\psi(tx, ty) = t^p \psi(x, y)$ of function $h_c := |y|^p - c|x|^p$ if $c < (p^* - 1)^p$.

Proof. Put $c_p = (p^* - 1)^p$. Fix (x^-, y^-) and (x^+, y^+) . Find z^- which almost gives supremum in $\varphi(x^-, y^-) = \sup [B(x^-, y^-, z) - c_p z]$. Do the same for $\varphi(x^+, y^+)$ to find z^+ . Then

$$B(x^-, y^-, z^-) - c_p z^- \leq \varphi(x^-, y^-) \leq B(x^-, y^-, z^-) - c_p z^- + \varepsilon,$$

$$B(x^+, y^+, z^+) - c_p z^+ \leq \varphi(x^+, y^+) \leq B(x^+, y^+, z^+) - c_p z^+ + \varepsilon.$$

Let $x = \frac{1}{2}(x^+ + x^-)$, $y = \frac{1}{2}(y^+ + y^-)$ and put $z = \frac{1}{2}(z^+ + z^-)$. Then

$$\varphi(x, y) = \sup \cdots \geq B(x, y, z) - c_p z = B(x, y, z) - c_p \frac{1}{2}(z^+ + z^-) \geq$$

$$\frac{1}{2}(B(x^-, y^-, z^-) - c_p z^-) + \frac{1}{2}(B(x^+, y^+, z^+) - c_p z^+) \geq \frac{1}{2}(\varphi(x^-, y^-) + \varphi(x^+, y^+)) - 2\varepsilon.$$

So φ is zigzag concave. Also

$$\varphi(x, y) = \sup \cdots \geq \lim_{z \rightarrow |x|^p +} [B(x, y, z) - c_p z] \geq |y|^p - c_p |x|^p = h(x, y).$$

So φ is a zigzag concave majorant of h . Why the least? Let ψ be any zigzag concave function such that

$$h \leq \psi.$$

Put $\Psi := \psi(x, y) + c_p z$. Then it is easy to see that Ψ satisfies (74), (75). Also on $\partial\Omega = \{z = |x|^p\}$ we have

$$\Psi(x, y, z) \geq h(x, y) + c_p z = h(x, y) + c_p |x|^p = |y|^p.$$

Then combination of the last inequality and the fact that Ψ satisfies (74), (75) gives (attention **exercise!**)

$$\Psi(x, y, z) \geq B(x, y, z).$$

This a non-trivial exercise. But then trivially for every (x, y)

$$\psi(x, y) = \sup_{z: (x, y, z) \in \Omega} [\Psi(x, y, z) - c_p z] \geq \sup_{z: (x, y, z) \in \Omega} [B(x, y, z) - c_p z] = \varphi(x, y).$$

We need now to prove that $h_c, c < c_p$ does not have zigzag concave homogeneous majorant.

This and more is done in

Lemma 3.8. Function $h_c, c < c_p$ does not have zigzag concave homogeneous majorant. If $c = c_p$, then the function $h_{c_p} =: h$ has such majorant given by

$$\Phi_0(x, y) := \begin{cases} |y|^p - (p^* - 1)^p = h(x, y), & \text{if } h \leq 0; \\ p \left(1 - \frac{1}{p^*}\right)^{p-1} (|y| + |x|)^{p-1} (|y| - (p^* - 1)|x|), & \text{if } h > 0. \end{cases}$$

Another zigzag concave majorant of $h = h_{c_p}$ (but not the least) is given by

$$\Phi(x, y) := p \left(1 - \frac{1}{p^*}\right)^{p-1} (|y| + |x|)^{p-1} (|y| - (p^* - 1)|x|).$$

Remark. The fact that $\Phi_0(x, y) \leq 0$ if $|x| \geq |y|$ will be crucial for the proof of Theorem 3.5.

The proof of Lemma 3.8. We work in the first quadrant. Homogeneous φ can be written as

$$\varphi(x, y) = (x+y)\varphi\left(\frac{x}{x+y}, \frac{y}{x+y}\right), \quad s := \frac{y-x}{y+x}, \quad \text{then } \frac{1+s}{2} = \frac{y}{x+y}, \quad \frac{1-s}{2} = \frac{x}{x+y}.$$

$$s'_x = -\frac{1+s}{x+y}, \quad s'_y = \frac{1-s}{x+y}.$$

We put $g(s) := \varphi(\frac{1-s}{2}, \frac{1+s}{2})$. Next we list some results of computations:

$$\varphi_x = p(x+y)^{p-1}g(s) + (x+y)^{p-1}g'(s)(1+s), \quad \varphi_y = p(x+y)^{p-1}g(s) + (x+y)^{p-1}g'(s)(1-s).$$

$$\varphi_{xx} = p(p-1)(x+y)^{p-2}g(s) - 2(p-1)(x+y)^{p-2}g'(s)(1+s) + (x+y)^{p-2}g''(s)(1+s)^2,$$

$$\varphi_{yy} = p(p-1)(x+y)^{p-2}g(s) + 2(p-1)(x+y)^{p-2}g'(s)(1-s) + (x+y)^{p-2}g''(s)(1-s)^2,$$

$$\varphi_{xy} = p(p-1)(x+y)^{p-2}g(s) - 2(p-1)(x+y)^{p-2}g'(s)s - (x+y)^{p-2}g''(s)(1-s^2).$$

So on $x+y=1$

$$\varphi_{xy} = -[(1-s^2)g''(s) + 2(p-1)sg'(s) - p(p-1)g(s)],$$

$$\varphi_{xx} + \varphi_{yy} = 2(1+s^2)g''(s) - 4(p-1)sg'(s) + 2p(p-1)g(s).$$

So combining the two:

$$(\partial_{x-y})^2\varphi = \varphi_{xx} - 2\varphi_{xy} + \varphi_{yy} = 4g''(s).$$

$$(\partial_{x+y})^2\varphi = \varphi_{xx} + 2\varphi_{xy} + \varphi_{yy} = 4g''(s) + 4\varphi_{xy} = 4(s^2g''(s) + (p-1)(-2sg'(s) + pg(s))).$$

Zigzag concave means the last two lines have sign ≤ 0 . To find φ satisfying these two ≤ 0 differential inequalities, let us try first to find it in such a way that the first inequality is equality! Hence we seek for the linear g ! Then put

$$(76) \quad g(s) = a \left(\frac{1+s}{2} - \rho \frac{1-s}{2} \right).$$

Then the second inequality $s^2g''(s) + (p-1)(-2sg'(s) + pg(s)) \leq 0$ becomes

$$(77) \quad 2sg'(s) - pg(s) \geq 0, \quad \text{on } [-1, 1].$$

It is satisfied (as g is linear) if and only if it is satisfied in -1 and 1 . We get

$$a(1 + \rho - p) \geq 0, \quad -a(1 + \rho + p\rho) \geq 0.$$

As g is greater than $\left(\frac{1+s}{2}\right)^p - c\left(\frac{1-s}{2}\right)^p$, it is positive at $s=1$, so $a > 0$. then we get from previous inequalities that

$$\rho \geq \max\left(p-1, \frac{1}{p-1}\right) = p^* - 1.$$

Let us try linear g with $\rho = p^* - 1$. So g has zero at s_p such that

$$p^* - 1 = \frac{1 + s_p}{1 - s_p}.$$

But if

$$h(x, y) = (x+y)^p H(s), \quad H(s) := \left(\frac{1+s}{2}\right)^p - c_p \left(\frac{1-s}{2}\right)^p,$$

then H has zero at the same point s_p . Now let us find a from the condition

$$H'(s_p) = g'(s_p) \Rightarrow a = p \left(1 - \frac{1}{p^*}\right)^{p-1}.$$

Where H is concave on $[-1, 1]$? Inflection point i_p is such that

$$\left(\frac{1+s}{2}\right)^{p-2} - c_p \left(\frac{1-s}{2}\right)^{p-2} = 0.$$

So it is clear that it is always $> s_p$. As H is concave on $[-1, i_p]$ it is concave on $[-1, s_p]$ (and a little bit on the right of s_p too).

It is also easy to see that on $[-1, s_p]$

$$(78) \quad s^2 H''(s) + (p-1)(-2sH'(s) + pH(s)) \leq 0.$$

This is an exercise.

Put now

$$\tilde{g}(s) = \begin{cases} \text{our linear } g(s) & s \in [s_p, 1]; \\ H(s), & s \in [-1, s_p]. \end{cases}$$

Then $\Phi_0(x, y) = (x+y)^p \tilde{g}(\frac{x}{x+y}, \frac{y}{x+y})$ is exactly the same Φ_0 as in Lemma 3.8's statement. We just checked that it is a zigzag concave majorant of $h(x, y)$. We also checked that $\Phi(x, y) = (x+y)^p g(\frac{x}{x+y}, \frac{y}{x+y})$, where g is our linear function built above, is zigzag concave majorant of $h(x, y)$ as well. It is exactly function Φ as in Lemma 3.8's statement.

Now let $c < c_p$. Linear function cannot be higher than corresponding H_c on $[-1, 1]$ and satisfy (77). In fact, if $\alpha s + \beta$ is higher, then $\alpha + \beta > 0$. Also (77) gives

$$2s\alpha - p(\alpha + \beta) \geq 0 \Rightarrow \alpha(2-p)s - p\beta \geq 0, \alpha(p-2) - p\beta \geq 0.$$

Then $\beta < 0, \alpha > 0$. So linear function is positive in 1 and negative at zero. So it must vanish on $[-1, 1]$, hence it has the form (76). Hence, we can see that their minorant H_c can have only $c \geq c_p$. In fact, we remember that $\rho \geq p^* - 1$ in (76). Then the zero of our linear function must be $\geq c_p$. But if our linear function is a majorant of H_c with $c < c_p$ it is also a majorant of H_{c_p} . Therefore, its zero must be $< c_p$. This is a contradiction, and a linear solution of two differential inequalities will not have minorant with $c < c_p$. Concave solution will not have such minorants either. **Exercise.**

Lemma 3.8 is finished. □

Theorem 3.7 is completely proved. □

Finishing the proof of Theorem 3.5. The real case. Now that we have function Φ (Φ_0 will work too) that is

1) zigzag concave on the plane,

2) is such that $\Phi(x, y) \geq h(x, y) := |y|^p - (p^* - 1)^p |x|^p$,

we can do the following. Fix f, g step functions on $I := [0, 1]$. Consider points $P = (x, y) = (\langle f \rangle_I, \langle g \rangle_I)$, $P^+ = (x^+, y^+) = (\langle f \rangle_{I_+}, \langle g \rangle_{I_+})$, $P^- = (x^-, y^-) = (\langle f \rangle_{I_-}, \langle g \rangle_{I_-})$. Notice that of course $P = \frac{1}{2}(P^+ + P^-)$. Also $|x^+ - x^-| = |y^+ - y^-|$ because this differences are $\frac{2}{\sqrt{|I|}}|(f, h_I)$ and $\frac{2}{\sqrt{|I|}}|(g, h_I)$ correspondingly, and we assumed in Theorem 3.5 that for every dyadic interval $|(f, h_I)| = |(g, h_I)|$. Let also $|x| \geq |y|$ (for example both are zeros)

Then we can use properties of Φ :

$$0 \geq \Phi(x, y) \geq \Phi(x^+, y^+)|I_+| + \Phi(x^-, y^-)|I_-|.$$

As intervals I^+, I^- are as good as I we can repeat this for them. Just iterating this procedure and denoting by I^σ dyadic intervals of size 2^{-n} with σ being any string of \pm of length n we get

$$\Sigma_\sigma \Phi(x^\sigma, y^\sigma)|I^\sigma| \leq 0.$$

Combine this with property 2) above. Then

$$\Sigma_\sigma |y^\sigma|^p |I^\sigma| \leq (p^* - 1)^p \Sigma_\sigma |x^\sigma|^p |I^\sigma|.$$

But by our construction $y^\sigma = \langle g \rangle_{I^\sigma}$, $x^\sigma = \langle f \rangle_{I^\sigma}$. So we get

$$\Sigma_\sigma |\langle g \rangle_{I^\sigma}|^p |I^\sigma| \leq (p^* - 1)^p \Sigma_\sigma |\langle f \rangle_{I^\sigma}|^p |I^\sigma|.$$

Going to the limit when $n \rightarrow \infty$ we get $\langle |g|^p \rangle_I \leq (p^* - 1)^p \langle |f|^p \rangle_I$, which gives the claim of Theorem 3.5 in the case of real-valued f, g .

□

Finishing the proof of Theorem 3.5. The complex-valued and Hilbert-valued cases. A certain “miracle” happens: Φ, Φ_0 have extra properties of symmetry, not apparent at this moment.

Extra symmetry. Consider $\varphi(x, y) = \Phi(\sqrt{x_1^2 + x_2^2}, \sqrt{y_1^2 + y_2^2})$. We use standard notations, now x, y are vectors, $\|\cdot\|$ is the norm of a vector, $dx := (dx_1, dx_2)$, $dy := (dy_1, dy_2)$ are also arbitrary vectors.

We want to see that

$$1) -d^2\varphi := -(H_\varphi \begin{bmatrix} dx \\ dy \end{bmatrix}, \begin{bmatrix} dx \\ dy \end{bmatrix}) \geq 0, \text{ if } \|dx\| = \|dy\|. \text{ This is “zigzag concavity”}$$

direct analog.

$$2) \varphi(x, y) \geq h(\|x\|, \|y\|).$$

The second is obvious, but the first happens by a “miracle”. Let us prove it and see, where the “miracle” happens.

Calculations (really abusing the language we understand that Φ_x, Φ_y are partial derivatives of Φ with respect to the first and the second variables):

$$\begin{aligned} \varphi_{x_1} &= \Phi_x \cdot \frac{x_1}{\sqrt{x_1^2 + x_2^2}}, \quad \varphi_{x_2} = \Phi_x \cdot \frac{x_2}{\sqrt{x_1^2 + x_2^2}}. \\ \varphi_{x_1 x_1} &= \Phi_{xx} \cdot \frac{x_1^2}{x_1^2 + x_2^2} + \Phi_x \frac{x_2^2}{(x_1^2 + x_2^2)^{3/2}}, \quad \varphi_{x_2 x_2} = \Phi_{xx} \cdot \frac{x_2^2}{x_1^2 + x_2^2} + \Phi_x \frac{x_1^2}{(x_1^2 + x_2^2)^{3/2}}. \\ \varphi_{x_1 x_2} &= \Phi_{xx} \cdot \frac{x_1 x_2}{x_1^2 + x_2^2} - \Phi_x \frac{x_1 x_2}{(x_1^2 + x_2^2)^{3/2}}. \end{aligned}$$

Symmetrically for y derivatives. Also

$$\varphi_{x_i y_j} = \Phi_{xy} \cdot \frac{x_i y_j}{\sqrt{x_1^2 + x_2^2} \sqrt{y_1^2 + y_2^2}}, \quad i, j = 1, 2.$$

Therefore,

$$\begin{aligned} -(H_\varphi \begin{bmatrix} dx \\ dy \end{bmatrix}, \begin{bmatrix} dx \\ dy \end{bmatrix}) &= \frac{\Phi_x}{\|x\|} \left(\frac{x_2 dx_1 - x_1 dx_2}{\|x\|} \right)^2 + \frac{\Phi_y}{\|y\|} \left(\frac{y_2 dy_1 - y_1 dy_2}{\|y\|} \right)^2 + \\ &\Phi_{xx} \left(\frac{x_2 dx_1 + x_1 dx_2}{\|x\|} \right)^2 + 2\varphi_{xy} \left(\frac{x_2 dx_1 + x_1 dx_2}{\|x\|} \right) \left(\frac{y_2 dy_1 + y_1 dy_2}{\|y\|} \right) + \varphi_{yy} \left(\frac{y_2 dy_1 + y_1 dy_2}{\|y\|} \right)^2 \end{aligned}$$

$$= \frac{\Phi_x}{\|x\|} \|\hat{dx}\|^2 + \frac{\Phi_y}{\|y\|} \|\hat{dy}\|^2 + \Phi_{xx}(dx, \frac{x}{\|x\|})^2 + 2\Phi_{xy}(dx, \frac{x}{\|x\|})(dy, \frac{y}{\|y\|}) + \Phi_{yy}(dy, \frac{y}{\|y\|})^2,$$

where \hat{dx}, \hat{dy} are projections of vectors dx, dy on direction orthogonal to x, y correspondingly.

Recall that up to a positive constant (which we drop now abusing the language)

$$\Phi(x, y) = (y - (p-1)x)(x+y)^{p-1}, \text{ if } p \geq 2,$$

and

$$(p-1)\Phi(x, y) = -(x - (p-1)y)(x+y)^{p-1}, \text{ if } p \leq 2.$$

Let us consider $p \geq 2$, the other case being similar. Looking at the formulae above we get by direct calculation with formula for Φ that for any numbers h', k'

$$\Phi_{xx}h'^2 + 2\Phi_{xy}h'k' + \Phi_{yy}k'^2 = -p(p-1)(x+y)^{p-2}(h'^2 - k'^2) - p(p-1)(p-2)x(x+y)^{p-3}(h' + k')^2.$$

(By the way we immediately see that this form is ≤ 0 if $|k'| = |h'|$, which is infinitesimal version of zigzag concavity.)

Now let us combine our formulae, putting $h' = (dx, \frac{x}{\|x\|}), k' = (dy, \frac{y}{\|y\|})$. Then

$$\begin{aligned} (H_\varphi \begin{bmatrix} dx \\ dy \end{bmatrix}, \begin{bmatrix} dx \\ dy \end{bmatrix}) &= \frac{\Phi_x}{\|x\|} \|\hat{dx}\|^2 + \frac{\Phi_y}{\|y\|} \|\hat{dy}\|^2 - p(p-1)(\|x\| + \|y\|)^{p-2}(h'^2 - k'^2) \\ &\quad - p(p-1)(p-2)\|x\|(\|x\| + \|y\|)^{p-3}(h' + k')^2. \end{aligned}$$

Let us look at the first line of the last formula. Calculate

$$\frac{\Phi_x}{\|x\|} \|\hat{dx}\|^2 + \frac{\Phi_y}{\|y\|} \|\hat{dy}\|^2 = \left(\frac{\Phi_x}{\|x\|} + \frac{\Phi_y}{\|y\|} \right) \|\hat{dy}\|^2 + \frac{\Phi_x}{\|x\|} (k'^2 - h'^2) + Term,$$

where $Term := \frac{\Phi_x}{\|x\|} (\|h\|^2 - \|k\|^2)$. This is just because $\|\hat{dx}\|^2 + h'^2 = \|h\|^2$, and the same is true for k . In particular,

$$(79) \quad Term = 0 \text{ if } \|h\| = \|k\|, \text{ and } Term \leq 0, \text{ if } \|h\| \geq \|k\|.$$

In fact,

$$(80) \quad \frac{\Phi_x}{\|x\|} = -p(p-1)(\|x\| + \|y\|)^{p-2} < 0.$$

Combine three last formulae. Then we have

$$\begin{aligned} (H_\varphi \begin{bmatrix} dx \\ dy \end{bmatrix}, \begin{bmatrix} dx \\ dy \end{bmatrix}) &= \left(\frac{\Phi_x}{\|x\|} + \frac{\Phi_y}{\|y\|} \right) \|\hat{dy}\|^2 \\ &\quad + Term - p(p-1)(p-2)\|x\|(\|x\| + \|y\|)^{p-3}(h' + k')^2. \end{aligned}$$

The second line is obviously negative (see (79)). To have the first line negative it is necessary and sufficient to have

$$(81) \quad \left(\frac{\Phi_x}{\|x\|} + \frac{\Phi_y}{\|y\|} \right) \leq 0.$$

Calculate:

$$\frac{\Phi_y}{\|y\|} = p(\|y\| - (p-2)\|x\|)(\|x\| + \|y\|)^{p-2}.$$

Combine this with (80) to get

$$\left(\frac{\Phi_x}{\|x\|} + \frac{\Phi_y}{\|y\|} \right) = -(\|x\| + \|y\|)^{p-2} \left(p(p-1) - p + p(p-2) \frac{\|x\|}{\|y\|} \right) =$$

$$(82) \quad -p(p-2) \frac{(\|x\| + \|y\|)^{p-1}}{\|y\|} \leq 0, \text{ if } p \geq 2.$$

The case $p < 2$ goes along the same lines with corresponding change in the formula for Φ . The proof of Theorem 3.5 is finished in the complex-valued case. One can notice that the same proof works in any Hilbert space, not just 2-dimensional as above, **exercise!**

□

Theorem is finally completely proved.

□

We want to remember a formula that has been just obtained ($p \geq 2$):

$$(83) \quad \begin{aligned} (H_\varphi \begin{bmatrix} dx \\ dy \end{bmatrix}, \begin{bmatrix} dx \\ dy \end{bmatrix}) &= -p(p-2) \frac{(\|x\| + \|y\|)^{p-1}}{\|y\|} \|\hat{d}y\|^2 - p(p-1)(\|x\| + \|y\|)^{p-2} (\|dx\|^2 - \|dy\|^2) \\ &\quad - p(p-1)(p-2)\|x\|(\|x\| + \|y\|)^{p-3} \left((dx, \frac{x}{\|x\|}) + (dy, \frac{y}{\|y\|}) \right)^2, \end{aligned}$$

where $\|\hat{d}y\|^2 = \|dy\|^2 - (dy, \frac{y}{\|y\|})^2$. Also

$$(84) \quad p \left(1 - \frac{1}{p^*} \right)^{p-1} \varphi \geq \|y\|^p - (p^* - 1)^p \|x\|^p.$$

On the other hand, if $1 < p < 2$, we know that

$$\varphi(x, y) = (\|y\|^p - (p^* - 1)\|x\|)(\|x\| + \|y\|)^{p-1}$$

will satisfy

$$(85) \quad \begin{aligned} (H_\varphi \begin{bmatrix} dx \\ dy \end{bmatrix}, \begin{bmatrix} dx \\ dy \end{bmatrix}) &= -p(2-p) \frac{(\|x\| + \|y\|)^{p-1}}{\|x\|} \|\hat{d}x\|^2 - p(p-1)(\|x\| + \|y\|)^{p-2} (\|dy\|^2 - \|dx\|^2) \\ &\quad - p(p-1)(2-p)\|x\|(\|x\| + \|y\|)^{p-3} \left((dx, \frac{x}{\|x\|}) + (dy, \frac{y}{\|y\|}) \right)^2, \end{aligned}$$

where $\|\hat{d}x\|^2 = \|dx\|^2 - (dx, \frac{x}{\|x\|})^2$. The same majorization (84) happens for $1 < p < 2$ as well.

Inequalities (52), (53) are completely done. However, to move further, in particular to (54), (55), we need a new tool=**stochastic integrals**.

4. STOCHASTIC INTEGRALS. ITÔ'S FORMULA

Let $w(s) := w_s$ denote Brownian motion started at 0, that is $w_0 = 0$, and for all $t_1 < t_2 < t_3$, random variables $w_{t_2} - w_{t_1}$, $w_{t_3} - w_{t_2}$ are Gaussian independent with zero average and variances $\sqrt{t_2 - t_1}$, $\sqrt{t_3 - t_2}$ correspondingly.

We want to understand what does it mean

$$\int_a^b \xi(t) dw_t.$$

It is **not** the Riemann sum definition as the following example shows.

Example. Consider two simplest Riemann sums built on a partition of the interval $[a, b]$:

$$\begin{aligned}\Sigma_1 &:= \sum_{i=1}^m w(t_{i-1})(w(t_i) - w(t_{i-1})), \\ \Sigma_2 &:= \sum_{i=1}^m w(t_i)(w(t_i) - w(t_{i-1})).\end{aligned}$$

If refinement is small, we should have had (if stochastic integral were a Riemann sum thing) that these two random variables Σ_1 and Σ_2 are close. Let us see, whether this is the case.

Notice that uniformly (when the partition changes) they are in $L^2(\Omega, \mathcal{F}, \mathbb{P})$, where $(\Omega, \mathcal{F}, \mathbb{P})$ is the probability space on which Brownian motion is given.

$$\begin{aligned}\mathbb{E}(\Sigma_1)^2 &= \sum_{i < j} \mathbb{E}(w(t_{i-1})(w(t_i) - w(t_{i-1}))w(t_{j-1}) \cdot (w(t_i) - w(t_{i-1}))) + \\ &\sum_i \mathbb{E}((w(t_{i-1})^2(w(t_i) - w(t_{i-1}))^2) = \sum_{i < j} \mathbb{E}(w(t_{i-1})(w(t_i) - w(t_{i-1}))w(t_{j-1}) \cdot \mathbb{E}(w(t_i) - w(t_{i-1})) + \\ &\sum_i t_i(t_i - t_{i-1}) \leq b(b - a).\end{aligned}$$

Also

$$\begin{aligned}\mathbb{E}(\Sigma_2)^2 &= 2\mathbb{E}\Sigma_1^2 + 2\mathbb{E}(\sum_i ((w(t_i) - w(t_{i-1})))^2)^2 \leq \\ &2b(b - a) + 2\mathbb{E}(\sum_i \xi_i^2)^2,\end{aligned}$$

where $\xi_i := w(t_i) - w(t_{i-1})$ are Gaussian independent with average zero and $\sigma_i^2 = |t_i - t_{i-1}|$. Then

$$\begin{aligned}\mathbb{E}(\sum_i \xi_i^2)^2 &= 2 \sum_{i < j} \mathbb{E}\xi_i^2 \mathbb{E}\xi_j^2 + \mathbb{E}\xi_i^4 = \\ &2 \sum_{i < j} |t_i - t_{i-1}| |t_j - t_{j-1}| + 3 \sum_i |t_i - t_{i-1}|^2 \leq 5(b - a)^2.\end{aligned}$$

The correct definition of integral should have been such that if Σ_1, Σ_2 are uniformly in L^2 and are both the Riemann sums, they should have been close in some sense. Suppose they are close (as random variables) in probability (one of the weakest sense possible). Then we use a simple **exercise** that if $\|f_n\|_{L^2(\mathbb{P})} \leq C$, $f_n \Rightarrow 0$, then $\|f_{n_k}\|_{L^1(\mathbb{P})} \rightarrow 0$.

In our case, nothing like that happened:

$$\begin{aligned}\mathbb{E}\Sigma_1 &= 0, \\ \mathbb{E}\Sigma_2 &= \mathbb{E}\Sigma_1 + \mathbb{E}(\sum_i (w(t_i) - w(t_{i-1}))^2) = \sum_i (t_i - t_{i-1}) = b - a \neq 0.\end{aligned}$$

We understand now that stochastic integral $\int_a^b \xi(t) dw(t)$ is a much more subtle thing than Riemann sum integral. Stochastic integrals were understood by Kioshi Itô.

4.1. A bit on Itô's definition. .

Let $B\mathcal{F}$ be a sigma algebra of sets $A \subset \mathbb{R} \times \Omega$ such that for every $t \in [a, b]$ we have $A \cap ((-\infty, t] \times \Omega)$ is in $B_t \times \mathcal{F}_t$, where B_t is Borel sigma algebra on $(-\infty, t]$, \mathcal{F}_t is a sigma algebra generated by $\{w_s\}_{s \leq t}$. Let $M_2[a, b]$ is the set of functions measurable with respect to $B\mathcal{F}$ such that

- (a) $f(t)$ is measurable with respect to \mathcal{F}_t for each t ,
- (b) with probability 1, $\int_a^b |f(t)|^2 dt < \infty$.

For all such random functions (random processes) Itô defines

$$(86) \quad \int_a^b f(t) dw(t).$$

Definition. $f \in M_2[a, b]$ is called a step function if there exists a partition such that $f(t) = f(t_i)(\omega)$, for $t \in [t_i, t_{i+1})$. We introduce the stochastic integral for them in a natural way

$$\int_a^b f dt := \sum_i f(t_i)(\omega) \cdot (w(t_{i+1}) - w(t_i)).$$

Lemma 4.1. *For every $f \in M_2[a, b]$ there exists a sequence of step functions as above such that with probability 1*

$$\lim_{n \rightarrow \infty} \int_a^b |f(t) - f_n(t)|^2 dt = 0.$$

Moreover if in addition

$$\mathbb{E} \int_a^b |f(t)|^2 dt < \infty,$$

then step functions can be chosen to have

$$\lim_{n \rightarrow \infty} \mathbb{E} \int_a^b |f(t) - f_n(t)|^2 dt = 0.$$

We need the following Lemma.

Lemma 4.2. *Let φ be a step function as above. Let $\delta, \epsilon > 0$. Then*

$$\mathbb{P}\left\{\left|\int_a^b \varphi(t) dw(t)\right| > \epsilon\right\} \leq \frac{\delta}{\epsilon^2} + \mathbb{P}\left\{\int_a^b |\varphi(t)|^2 dt > \delta\right\}.$$

This lemma immediately gives the following reasoning. If—as above— $f \in M_2[a, b]$ and f_n are step functions from Lemma 4.1, then

$$\mathbb{P} - \lim_{n \rightarrow \infty} \int_a^b |f(t) - f_n(t)|^2 dt = 0.$$

Then

$$\mathbb{P} - \lim_{n, m \rightarrow \infty} \int_a^b |f_m(t) - f_n(t)|^2 dt = 0.$$

By definition

$$\forall \epsilon > 0, \mathbb{P}\left\{\int_a^b |f_m(t) - f_n(t)|^2 dt > \epsilon\right\} \rightarrow 0, m, n \rightarrow \infty.$$

Now we use Lemma 4.2 to have

$$\limsup_{m,n \rightarrow \infty} \mathbb{P}\{|\int_a^b f_n(t) dw(t) - \int_a^b f_m(t) dw(t)| > \varepsilon\} \leq \frac{\delta}{\varepsilon^2}$$

for any $\delta > 0$. So the sequence of random variables $\xi_n := \int_a^b f_n(t) dw(t)$ is Cauchy convergent in measure (in probability). So in probability it converges to a certain random variable ξ . This ξ is **by definition** $\int_a^b f dw(t)$. **Itô's stochastic integral is constructed.**

This integral has many nice properties:

If $\mathbb{E} \int_a^b |f(t)|^2 dt < \infty$, then $\mathbb{E} \int_a^b f(t) dw(t) = 0$ and

$$\mathbb{E}(\int_a^b f(t) dw(t))^2 = \mathbb{E} \int_a^b |f(t)|^2 dt.$$

If in addition $\mathbb{E} \int_a^b |g(t)|^2 dt < \infty$ then

$$(87) \quad \mathbb{E}(\int_a^b f(t) dw(t) \cdot \int_a^b g(t) dw(t)) = \mathbb{E} \int_a^b f(t) \cdot g(t) dt.$$

(Integral of the product is the product of integrals.)

4.2. Stochastic differential. Let $b(t) \in M_2[a, b]$, and $a(t)$ be measurable with respect to \mathcal{F}_t for every t , and

$$\int_a^b |a(t)| dt < \infty.$$

Suppose $\zeta(t)$ is a random process such that for all t_1, t_2 such that $a \leq t_1 \leq t_2 \leq b$

$$\zeta(t_2) - \zeta(t_1) = \int_{t_1}^{t_2} a(t) dt + \int_{t_1}^{t_2} b(t) dw(t).$$

Then we write the above line as **stochastic differential**:

$$d\zeta(t) = a(t)dt + b(t)dw(t).$$

Remark. If $a = 0$ this integral is a martingale (obviously) on the filtration $\{\mathcal{F}_t\}_{t>0}$ of sigma algebras generated by Brownian motions.

4.3. Itô' formula. Let ζ have the stochastic differential in the sense above and let $u(t, x)$ be a (several times) smooth function. Consider new process

$$\eta(t) := u(t, \zeta(t)).$$

Theorem 4.3. *Then η also has stochastic differential and*

$$d\eta(t) = [u'_t(t, \zeta(t)) + u'_x(t, \zeta(t))a(t) + \frac{1}{2}u''_{xx}(t, \zeta(t)) \cdot b^2(t)] dt + u'_x(t, \zeta(t)) \cdot b(t) \cdot dw(t).$$

Proof. The proof is quite subtle. See [45], [62].

□

Matrix Itô's formula also exists and will be used. Let a be $m \times 1$ column of processes, σ is a $m \times k$ matrix of processes (with entries in $M_2[a, b]$). Let $W(t)$ be a column of k independent Brownian motion. Let $\zeta(t)$ be a $m \times 1$ process with stochastic differential

$$d\zeta(t) = a(t) dt + \sigma dW(t).$$

Let $u(t, x)$ be a smooth function, where $x \in \mathbb{R}^m$. Let $\eta(t) = u(t, \zeta(t))$. Then η also has stochastic differential, and matrix Itô's formula gives

$$(88) \quad d\eta(t) = [\partial u / \partial t + \nabla_x u(t, \zeta) \cdot a(t) + \frac{1}{2} \text{trace}(\sigma H_u(t, \zeta) \sigma^*)] dt + \nabla_x u \cdot \sigma dW(t).$$

Here \cdot is the scalar product in \mathbb{R}^m .

4.4. Space-time Brownian motion. Let us discuss Theorem 4.3. If $a = 0$, then the process ζ is a martingale (see Remark before the theorem). However, it is quite unrealistic to expect that if we consider the composition of a non-linear function u and a martingale, then we would get another martingale. And in fact, if $a = 0$ the formula in Theorem 4.3 becomes (if $a = 0$)

$$(89) \quad d\eta(t) = [u'_t(t, \zeta(t)) + \frac{1}{2} u''_{xx}(t, \zeta(t)) \cdot b^2(t)] dt + u'_x(t, \zeta(t)) \cdot b(t) \cdot dw(t),$$

and the “non-martingale” part (called **drift**) in square brackets is very much present. But there is one very important exception.

Suppose $f \in C_0^\infty$ and $u^f(t, x)$ is the heat extension of f , in other words, the solution of the heat equation:

$$(90) \quad \left(\frac{\partial}{\partial t} - \frac{1}{2} \frac{\partial^2}{\partial x^2} \right) u^f = 0, \quad u^f(0, x) = f(x).$$

Fix large positive T and consider function of (t, x) given by $u = u^f(T - t, x)$. We want to compose it with stochastic process as in Theorem 4.3, with $a = 0, b = 1$. Then we get the process

$$\eta := u^f(T - t, w_t).$$

It will be a martingale on $[0, T]$. In fact, we can use (89) to get

$$d\eta(t) = \left[-\frac{\partial u^f}{\partial t}(T - t, w_t) + \frac{1}{2} \frac{\partial^2 u^f}{\partial x^2}(T - t, w_t) \right] dt + \frac{\partial u^f}{\partial x}(T - t, w_t) dw_t,$$

and by (90) the drift term in the brackets disappears.

If we work with heat extension for functions on \mathbb{R}^k the same will be true. Now Brownian motion W_t is k -dimensional (just k independent Brownian motions) and u^f is the solution of heat equation

$$(91) \quad \left(\frac{\partial}{\partial t} - \frac{1}{2} \Delta \right) u^f = 0, \quad u^f(0, x) = f(x).$$

Then we get the following stochastic differential

$$(92) \quad du^f(T - t, W_t) = \nabla_x u^f(T - t, W_t) \cdot dW_t,$$

where \cdot is the scalar product in \mathbb{R}^k .

We are interested now in the case of complex valued function f on \mathbb{R}^2 , so $k = 2$. Thinking that gradient is always column vector and W_t is 2-dimensional row vector it is convenient to rewrite (92) as

$$(93) \quad f(W_T) - u^f(T, 0) = \int_0^T dW_t \cdot \begin{bmatrix} \partial_x \\ \partial_y \end{bmatrix} u^f(T - t, W_t).$$

Definition. The expressions $\int_0^T dW_t \cdot \begin{bmatrix} \partial_x \\ \partial_y \end{bmatrix} u^f(T - t, W_t)$ will be called **heat martingales**.

But we will need a bigger class, where heat martingales are supplemented by their **martingale transforms**. The simplest martingale transforms are given by expressions

$$\int_0^T dW_t \cdot A \begin{bmatrix} \partial_x \\ \partial_y \end{bmatrix} u^f(T - t, W_t),$$

where A is a fixed matrix not depending neither on ω (elementary event) nor on time t .

Consider a special matrix

$$(94) \quad A := \begin{bmatrix} 1, & i \\ i, & -1 \end{bmatrix}$$

Then we get

$$\int_0^T dW_t \cdot \begin{bmatrix} \partial_x + i\partial_y \\ i(\partial_x + i\partial_y) \end{bmatrix} u^f(T - t, W_t),$$

which is

$$(95) \quad 2 \int_0^T dW_t \cdot \begin{bmatrix} \bar{\partial} \\ i\bar{\partial} \end{bmatrix} u^f(T - t, W_t).$$

This is quite suggestive. In fact, denoting temporarily the Ahlfors–Beurling transform $R_1^2 - R_2^2 + 2iR_1R_2$ by symbol AB , we recall that $AB\bar{\partial} = \partial$. The following theorem holds.

Theorem 4.4.

$$1) \quad \lim_{T \rightarrow \infty} \mathbb{E} \left(\int_0^T dW_t \cdot \begin{bmatrix} \partial_x \\ \partial_y \end{bmatrix} u^f(T - t, W_t) | W_T = z \right) = f(z),$$

$$2) \quad \lim_{T \rightarrow \infty} \mathbb{E} \left(\int_0^T dW_t \cdot \begin{bmatrix} \partial_x + i\partial_y \\ i(\partial_x + i\partial_y) \end{bmatrix} u^f(T - t, W_t) | W_T = z \right) = AB(f)(z).$$

Proof. Let us consider a test function g and build a heat martingale $X(t)$, $0 \leq t \leq T$, by formula 1), but with f replaced by g :

$$X(t) := g(T, 0) + \int_0^t dW_s \cdot \begin{bmatrix} \partial_x \\ \partial_y \end{bmatrix} u^g(T - s, W_s).$$

Let $Y(t)$, $0 \leq t \leq T$, denote the martingale in formula 2):

$$Y(t) := \int_0^t dW_s \cdot \begin{bmatrix} \partial_x + i\partial_y \\ i(\partial_x + i\partial_y) \end{bmatrix} u^f(T - s, W_s).$$

Then by “rule” that the product of stochastic integrals is “the integral of the product”, we get (below $k(t; x, y) := \frac{1}{2\pi t} e^{-\frac{x^2+y^2}{t}}$)

$$\begin{aligned} 2\pi T \mathbb{E}(Y(T) \cdot X(T)) &= 2\pi T \int_0^T \iint_{\mathbb{R}^2} \bar{\partial} u^f(T-t; x, y) \bar{\partial} u^g(T-t; x, y) k(t; x, y) dx dy dt \\ &= -2\pi T \int_0^T \iint_{\mathbb{R}^2} \bar{\partial} u^f(t; x, y) \bar{\partial} u^g(t; x, y) k(T-t; x, y) dx dy dt. \end{aligned}$$

Notice that $2\pi T k(T-t; x, y) \rightarrow 1$ if T goes to infinity. It is not then difficult to see that the last expression becomes very close to

$$\int_0^T \iint_{\mathbb{R}^2} \bar{\partial} u^f(t; x, y) \bar{\partial} u^g(t; x, y) dx dy dt,$$

when T goes to infinity.

Recall formula (30) and formula $AB = R_1^2 - R_2^2 + 2iR_1R_2$. (Number 2 in (30) should be dropped now as we are working with extensions with respect to $\frac{\partial}{\partial t} - \frac{1}{2}\Delta$ unlike before formula (30), where we worked with $\frac{\partial}{\partial t} - \Delta$.) Combined they give us that the last expression would be equal to $(AB(f), g)$ if the integration would be $\int_0^\infty \dots dt$ and not $\int_0^T \dots dt$. But as T is large and f, g are nice the “error” goes to zero when T goes to infinity. So

$$(96) \quad 2\pi T \mathbb{E}(Y(T) \cdot X(T)) = (AB(f), g) + o(1).$$

On the other hand, $X(T) = g(W_T)$ by (93) with f replaced by g . Therefore, for any test function g

$$2\pi T \mathbb{E}(Y(T) \cdot g(W_T)) = 2\pi T \int_{\mathbb{C}} d\mu_T(z) \mathbb{E}(Y(T) | W_T = z) g(z),$$

where $d\mu_T = \frac{1}{2\pi T} e^{-\frac{|z|^2}{T}} dm_2(z)$ is given by the density distribution of W_T . Now using the facts that g is a nice test function and that $2\pi T \frac{d\mu_T(z)}{dm_2(z)} \rightarrow 1$ pointwise and in a bounded fashion when $T \rightarrow \infty$ we obtain

$$\int_{\mathbb{C}} \mathbb{E}(Y(T) | W_T = z) g(z) dm_2(z) = 2\pi T \mathbb{E}(Y(T) \cdot g(W_T)) + o(1).$$

Comparing this with (96) we get the formula

$$(97) \quad AB(f)(z) = \lim_{T \rightarrow \infty} \mathbb{E} \left(\int_0^T dW_t \cdot A \nabla_{x,y} u^f(T-t; W_t) | W_T = z \right).$$

Theorem is proved. □

Remark. It is very easy to see now that for martingale $\{Y(t)\}_{0 \leq t \leq T}$ constructed above

$$\|AB(f)\|_{L^p(\mathbb{C}, dm_2)}^p \leq \lim_{T \rightarrow \infty} 2\pi T \mathbb{E}|Y(T)|^p$$

for any p . It is a sort of averaging operator. Moreover, for martingale $\{X(t)\}_{0 \leq t \leq T}$ we obviously have limiting equality

$$\|g\|_{L^p(\mathbb{C}, dm_2)}^p = \lim_{T \rightarrow \infty} 2\pi T \mathbb{E}|X(T)|^p.$$

This is trivial from (93): just raise both part to the power p and take the expectation (first conditioning over $W_T = z$, then integrating with respect to $d\mu_T(z)$) and use again the fact that $2\pi T \frac{d\mu_T(z)}{dm_2(z)} \rightarrow 1$.

Now put $g = f$. We see that

$$\|AB(f)\|_p \leq 2M_p\|f\|_p$$

follows from $\mathbb{E}|Y|^p \leq M_p\mathbb{E}|X|^p$ for martingales Y, X . Notice that Y is just the martingale transform of X with the help of matrix A , whose norm is 2. This explains the constant 2 in the above display inequality. This is why we study below X, Y and their relationship.

Remark. The reader can find many interesting examples, references and explanations in recent review of Banuelos devoted to Burkholder's estimate: [4].

4.5. Orthogonal (conformal) martingales. Introducing two martingales on the filtration of Brownian motion

$$X(t) := \int_0^t dW_s \cdot \nabla_{x,y} u^f(T-s; W_s), 0 \leq t \leq T;$$

$$Y(t) := \int_0^t dW_s \cdot A \nabla_{x,y} u^f(T-s; W_s) =: A \star X(t), 0 \leq t \leq T,$$

and using the previous remark, we get that it might be a fruitful idea to look for a sharp **martingale transform inequality**

$$(98) \quad \mathbb{E}\|A \star X(t)\|_p \leq M_p \mathbb{E}\|X\|_p.$$

Let $f = \phi + i\psi$. Introduce the notations

$$H_1^s = [u_x^\phi(T-s; W_s), u_x^\psi(T-s; W_s)] ,$$

$$H_2^s = [v_y^\phi(T-s; W_s), v_y^\psi(T-s; W_s)] .$$

$$K_1^s = [u_x^\phi(T-s; W_s) - u_y^\psi(T-s; W_s), u_y^\phi(T-s; W_s) + u_x^\psi(T-s; W_s)] ,$$

$$K_2^s = [-u_y^\phi(T-s; W_s) - u_x^\psi(T-s; W_s), u_x^\phi(T-s; W_s) - u_y^\psi(T-s; W_s)] ,$$

we can write complex martingale $X = X_1 + iX_2$, $Y = A \star X = Y_1 + iY_2$ in the form (below dW_s is a 2-row-vector)

$$(99) \quad [X_1(t), X_2(t)] = \int_0^t dW_s \begin{bmatrix} H_1^s \\ H_2^s \end{bmatrix} = \int_0^t (H_1^s dw_s^1 + H_2^s dw_s^2) .$$

$$(100) \quad [Y_1(t), Y_2(t)] = \int_0^t dW_s \begin{bmatrix} K_1^s \\ K_2^s \end{bmatrix} = \int_0^t (K_1^s dw_s^1 + K_2^s dw_s^2) .$$

Properties of H, K . Vector processes H and K are related by

$$(101) \quad \|K_1^s\|^2 + \|K_2^s\|^2 \leq 4(\|H_1^s\|^2 + \|H_2^s\|^2)$$

for all elementary events ω and all times s .

Relationship (101) is called **differential subordination** of martingale Y to martingale $2X$.

Theorem 4.5 (Burkholder's theorem). *If martingale M is differentially subordinated to martingale N , then*

$$\mathbb{E}\|M(t)\|_p \leq (p^* - 1)\mathbb{E}\|N(t)\|_p, p^* := \max(p, p/p - 1) .$$

The constant is sharp.

In particular,

$$(102) \quad \mathbb{E}\|A \star X\|_p \leq 2(p^* - 1)\mathbb{E}\|X(t)\|_p.$$

But the constant is not sharp! We followed the probabilistic proof in [7], which “randomize” the idea of [58]. There is an analytic proof following [58] more directly, see [49].

More properties of H, K . Vector processes K have extra properties:

$$(103) \quad K_1^s \cdot K_2^s = 0, \quad \|K_1^s\| = \|K_2^s\|$$

for all elementary events ω and all times s . Such martingales are called **orthogonal or conformal**.

Theorem 4.6 (Banuelos–Janakiraman’s theorem). *We make the exposition of [5] using the notations above. If martingale M is differentially subordinated to martingale N , and martingale M is conformal and $\|M(0)\| \leq \|N(0)\|$ then*

$$\mathbb{E}\|M(t)\|_p \leq \sqrt{\frac{p(p-1)}{2}} \mathbb{E}\|N(t)\|_p, \quad p > 2.$$

In particular,

$$(104) \quad \mathbb{E}\|A \star X\|_p \leq \sqrt{2p(p-1)} \mathbb{E}\|X(t)\|_p, \quad p > 2.$$

But the constant is not sharp!

However, this inequality gives

$$\|T\|_p \leq \sqrt{2p(p-1)}, \quad p > 2.$$

Interpolation between $p = 2$ and large p with this estimate, optimization in this large p , will give (54).

The proof of Theorem 4.6. Our main tool will be formula (83). Trivial renormalization shows that to prove Theorem 4.6 it is enough to prove that if M, N are two martingales on the filtration of 2-dimensional Brownian motion and M is differentially subordinated to $\sqrt{\frac{2(p-1)}{p}} \cdot N$, $p > 2$, and M is **conformal** then

$$(105) \quad \mathbb{E}\|M(t)\|_p \leq (p-1)\mathbb{E}\|N(t)\|_p, \quad p > 2.$$

Consider such M, N , and their H_1, H_2, K_1, K_2 . We know that

$$(106) \quad \|K\|^2 \leq \frac{2(p-1)}{p} \|H\|^2,$$

where $\|K\|^2 := \|K_1\|^2 + \|K_2\|^2$, $\|H\|^2 := \|H_1\|^2 + \|H_2\|^2$, and

$$k_{11} \cdot k_{21} + k_{12} \cdot k_{22} = 0.$$

This and equality $\|K_1\| = \|K_2\|$ easily implies

$$(107) \quad k_{11} \cdot k_{12} + k_{21} \cdot k_{22} = 0.$$

Let $V(M, N) := \|M\|^p - (p-1)\|N\|^p$, $p > 2$, $\varphi(M, N) := p(1 - 1/p)^{p-1}(\|M\| + \|N\|)^{p-1}(\|M\| - (p-1)\|N\|)$.

We would like to prove that $\mathbb{E}(V(M(t), N(t))) \leq 0$. But it has been proved that $V \leq \varphi$. So it is enough to prove

$$(108) \quad \mathbb{E}(\varphi(M(t), N(t))) \leq 0.$$

To prove (108) we use:

$$(109) \quad \varphi(M(t), N(t)) = \varphi(M(0), N(0)) + \int_0^t d\varphi(M(s), N(s)).$$

To compute $\mathbb{E}d\varphi$ we use Itô's formula, which of course involves Hessian H_φ . More precisely, $d\varphi(s)$ will involve

$$(H_\varphi \begin{bmatrix} H_1^s \\ K_1^s \end{bmatrix}, \begin{bmatrix} H_1^s \\ K_1^s \end{bmatrix}) + (H_\varphi \begin{bmatrix} H_2^s \\ K_2^s \end{bmatrix}, \begin{bmatrix} H_2^s \\ K_2^s \end{bmatrix})$$

Now we look at formula (83), which gives

$$d\varphi = p(1-1/p)^{p-1}(A+B+C+D), \quad A := -p(p-1)(\|M(s)\| + \|N(s)\|)^{p-2}(\|H\|_2^2 - \|K\|_2^2),$$

$$B := -p(p-2)(\|M(s)\| + \|N(s)\|)^{p-1}\|M(s)\|^{-1} \left[\left(\frac{M_2 k_{11} - M_1 k_{12}}{\|M\|} \right)^2 + \left(\frac{M_2 k_{21} - M_1 k_{22}}{\|M\|} \right)^2 \right],$$

where we need to recall that $M = M_1 + iM_2$, M_1, M_2 being its real and imaginary parts. Part C comes from the last part of formula (83), and, obviously,

$$C \leq 0,$$

$$D = \dots dw_s^1 + \dots dw_s^2,$$

where \dots involve functions of k_{ij}^s, h_{ij}^s and $\nabla\varphi(M(s), N(s))$. This shows that $\int_0^t D(s)$ is a martingale starting at 0 and so

$$(110) \quad \mathbb{E} \int_0^t D(s) = 0.$$

We open the brackets in B , use (107), and the fact that $k_{11}^2 + k_{21}^2 = k_{12}^2 + k_{22}^2$, to get

$$B \leq -p(p-2)\left(\frac{1}{2}\|K\|_2^2\right)(\|M\| + \|N\|)^{p-2}.$$

Now

$$A + B = -p(\|M\| + \|N\|)^{p-2}[(p-1)\|H\|_2^2 - \frac{p}{2}\|K\|_2^2] \leq 0,$$

if (106) is valid. Term C is non-positive. Term D disappears after integration $\mathbb{E} \int_0^t$, As a result we come to (see (109)):

$$\mathbb{E} \int_0^t d\varphi(M(s), N(s)) = \mathbb{E}\varphi(M(0), N(0)) = \varphi(M(0), N(0)) \leq 0,$$

because if $x := \|M(0)\| \leq y := \|N(0)\|$ then $\varphi(x, y) \leq 0$, which is obvious from the formula for φ . □

As we already mentioned, this proves (54). To prove (55) one needs even more careful stochastic analysis, and we leave this for the next round of lectures somewhere in the future.

5. BELLMAN FUNCTION OF STOCHASTIC OPTIMAL CONTROL PROBLEMS

Let W_s be d_1 dimensional Brownian motion. Let $x(t)$ is a d -dimensional process given by

$$(111) \quad x(t) = x + \int_0^t b(\alpha(s), x(s)) ds + \int_0^t \sigma(\alpha(s), x(s)) dW_s,$$

in other words the process starts at $x \in \mathbb{R}^d$ and satisfies a stochastic differential equation

$$dx(t) = b(\alpha(t), x(t)) dt + \sigma(\alpha(t), x(t)) dW_t,$$

where α is a d_2 -dimensional control process, we can choose it ourselves, but it must be adapted, that is $\alpha(s)$ has to be measurable with respect to sigma algebra \mathcal{F}_s generated by $W_t, 0 \leq t \leq s$. Also values of the process α are often restricted: $\alpha(s, \omega) \in A \subset \mathbb{R}^{d_2}$.

Matrix function σ is smooth and $d \times d_1$ -dimensional, and b is a smooth column function of size d . Everything happens in $\Omega \subset \mathbb{R}^d$ (often $= \mathbb{R}^d$).

The choice of adapted process $\alpha(s)$ gives us different motions, all started at the same initial $x \in \mathbb{R}^d$.

This is a “broom” of motions, hidden elementary even ω gives “one stem of a broom”.

Suppose we are given the profit function $f(\alpha, x)$, meaning that on a trajectory of $x(t)$, for the time interval $[t, t + \Delta t]$, the profit is $f(\alpha(t), x(t)) + o(\Delta t)$. So on the whole trajectory we earn

$$\int_0^\infty f(\alpha(t), x(t)) dt.$$

We are also given the pension—we call it bonus function F —how much one is given at the end of the life. We want to choose a control process $\alpha = \alpha(s)$ to **maximize the average profit**:

$$(112) \quad v^\alpha(x) := \mathbb{E} \int_0^\infty f(\alpha(t), x(t)) dt + \limsup_{t \rightarrow \infty} \mathbb{E} F(x(t)).$$

If $b = 0$ and F is convex then one can write \lim instead of \limsup .

The optimal average gain, or

$$(113) \quad v := \sup_{\alpha} v^\alpha(x)$$

is called the Bellman function of stochastic optimal control problem (111), (112).

Usually the analysis consists of

a) writing Bellman PDE on v ;

b) solving it;

c) using “verification theorem”, which says that under certain conditions on data $\sigma, b, F, f, \Omega, A$ the classical solution of Bellman PDE is exactly v from (113).

5.1. Writing Bellman PDE.. This consists of a) Itô's formula, b) Bellman's principle of dynamic programming.

Using Itô's formula (88) we get

$$\begin{aligned} dv(x(s)) &= \sum_{k=1}^d \frac{\partial v}{\partial x_k}(x(s)) \sum_{j=1}^{d_1} \sigma_{kj}(\alpha(s), x(s)) dw_s^j + \\ &\quad \sum_{k=1}^d \frac{\partial v}{\partial x_k}(x(s)) b_k(\alpha(s), x(s)) ds + \\ &\quad \frac{1}{2} \sum_{i,j=1}^d \frac{\partial^2 v}{\partial x_i \partial x_j}(x(s)) a^{ij}(\alpha(s), x(s)), \end{aligned}$$

where

$$a^{ij}(\alpha, x) := \sum_{k=1}^{d_1} \sigma_{ik}(\alpha, x) \sigma_{kj}(\alpha, x)$$

is i, j matrix element of $d \times d$ matrix $\sigma \sigma^*$.

Introduce two linear differential operators with non-constant coefficients:

$$\begin{aligned} \mathcal{L}_1(\alpha, x) &:= \sum_{k=1}^d b_k(\alpha, x) \frac{\partial}{\partial x_k}, \\ \mathcal{L}_2(\alpha, x) &:= \sum_{i,j=1}^d a_{ij}(\alpha, x) \frac{\partial^2}{\partial x_i \partial x_j}, \end{aligned}$$

and

$$\mathcal{L}(\alpha, x) := \mathcal{L}_1(\alpha, x) + \mathcal{L}_2(\alpha, x).$$

Let us hit our formula for $dv(x(t))$ above by the expectation \mathbb{E} , then the first line becomes 0, and we get

$$\mathbb{E} \left[\frac{d}{dt} v(x(t)) \right] = \mathbb{E}[\mathcal{L}(\alpha(t), x(t))v](x(t)).$$

Or

$$(114) \quad \mathbb{E}v(x(t)) = v(x) + \mathbb{E} \int_0^t [\mathcal{L}_1(\alpha(s), x(s)) + \mathcal{L}_2(\alpha(s), x(s))]v(x(s)) ds.$$

Now we need the second ingredient to write the Bellman equation: **the Bellman principle= dynamic programming principle**. It is in this next equality:

$$\begin{aligned} v(x) &= \sup_{\alpha} \mathbb{E} \left[\int_0^{\infty} f(\alpha(t), x(t)) dt + \limsup_{t \rightarrow \infty} \dots \right] \\ (115) \quad &= \sup_{\alpha} \mathbb{E} \left[\int_0^t f(\alpha(t), x(t)) dt + v(x(t)) \right], \quad \forall t > 0. \end{aligned}$$

A minute though shows that this reflects the stationarity of Brownian motion and the fact that to be perfect one has to be perfect every second.

Now plug $\mathbb{E}v(x(t))$ from (114) into (115). We get

$$0 = \sup_{\alpha} \mathbb{E} \left[\int_0^{\infty} f(\alpha(t), x(t)) + \mathcal{L}(\alpha(t), x(t))v(x(t)) dt \right], \quad \forall t > 0.$$

Divide by t and tend t to zero. We “obtain” **Bellman equation**:

$$(116) \quad \sup_{\alpha \in A} [(\mathcal{L}(\alpha, x)v)(x) + f(\alpha, x)] = 0.$$

Positivity (usually present) of f and convexity (usually present) of F imply (if there is no drift, that is if $b(\alpha, x) = 0$) **obstacle condition**:

$$(117) \quad v(x) \geq F(x), \quad \forall x \in \Omega.$$

Often it becomes **boundary condition**:

$$(118) \quad v(x) = F(x), \quad \forall x \in \partial\Omega.$$

The definition of v in domain (not in the whole \mathbb{R}^d) should be slightly changed. The integration of profit function now is not from zero to infinity, but from zero to the stopping time of the first hit of $\partial\Omega$ by the trajectory $x(t)$.

See details of obtaining (116) in the beautiful book of N. Krylov [47].

In applications one is also interested in supersolutions of the Bellman equation (116) :

$$(119) \quad \begin{cases} \sup_{\alpha \in A} [\mathcal{L}(\alpha, x)V(x) + f(\alpha, x)] \leq 0, & x \in \Omega, \\ V(x) \geq F(x), & x \in \Omega. \end{cases}$$

Lemma : Let V solves (4) and let v be the Bellman function, then $V \geq v$ in Ω .

Proof : Equation (116) states that $-\mathcal{L}(\alpha, x)V(x) \geq f(\alpha, x)$. Using (114) for V and then (119), one gets

$$\begin{aligned} V(x) &= \mathbb{E}V(x(t)) - \mathbb{E} \int_0^t (\mathcal{L}(\alpha(s), x(s))V)(x(s)) ds \\ &\geq \mathbb{E}F(x(t)) + \mathbb{E} \int_0^t f(\alpha(s), x(s)) ds. \end{aligned}$$

Writing $\overline{\lim}_{t \rightarrow \infty}$ of both parts, we get $V(x) \geq v^\alpha(x)$. It rests to take the supremum over the control process α .

5.2. Special matrices σ bring us to Harmonic Analysis. Let us consider a very simple matrix σ not depending on x :

$$(120) \quad d_1 = 1, \quad \sigma(\alpha, x) = \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_d \end{bmatrix} =: \alpha.$$

If on the top of that $b = 0$ then operator \mathcal{L} just involves Hessian matrix H_v of function v :

$$(\mathcal{L}(\alpha)v)(x) = \frac{1}{2} \sum_{i,j=1}^d \frac{\partial^2}{\partial x_i \partial x_j} v(x) = \frac{1}{2} (H_v(x)\alpha, \alpha).$$

We claim that this is the generic case of Harmonic Analysis problems in \mathbb{R}^1 . Equation (116) becomes

$$(121) \quad \begin{cases} \sup_{\alpha \in A} [\frac{1}{2} \langle H_v(x) \alpha, \alpha \rangle + f(\alpha, x)] = 0, x \in \Omega, \\ v(x) \geq F(x), x \in \Omega. \end{cases}$$

If $b \neq 0$ then we just add the first order differential operator (called **drift**):

$$(122) \quad \begin{cases} \sup_{\alpha \in A} [\frac{1}{2} \langle H_v(x) \alpha, \alpha \rangle + \sum_{k=1}^d b_k(\alpha, x) \frac{\partial}{\partial x_k} v(x) + f(\alpha, x)] = 0, x \in \Omega, \\ v(x) \geq F(x), x \in \Omega. \end{cases}$$

Harmonic analysis on \mathbb{R}^2 “becomes” the analysis of the following Bellman equation (and this is exactly what we did in the sections above devoted to the analysis of the Ahlfors–Beurling operator):

$$(123) \quad d_1 = 2, \sigma(\alpha, x) = \begin{bmatrix} \alpha_{11} & \alpha_{12} \\ \vdots & \vdots \\ \alpha_{d1} & \alpha_{d2} \end{bmatrix} =: \alpha.$$

Conformal restrictions: Matrix α can have restrictions $\alpha \in A$ of the type that the first row is orthogonal to the second row and that the norms of the rows are equal. The reader can notice that these are **Cauchy–Riemann** conditions, and the corresponding solution of (111) will be a conformal martingale (again if $b = 0$). Bellman equation becomes

$$(124) \quad \begin{cases} \sup_{\alpha \in A} [\frac{1}{2} \text{trace}(\alpha^* H_v(x) \alpha) + \sum_{k=1}^d b_k(\alpha, x) \frac{\partial}{\partial x_k} v(x) + f(\alpha, x)] = 0, x \in \Omega, \\ v(x) \geq F(x), x \in \Omega. \end{cases}$$

Remarks. 1) This is (exactly as (122)) very non-linear (actually an example of so-called **fully non-linear**) equation of the second order.

2) This equation is much more difficult to analyze than (122). On the other hand, we can easily notice that conformal restrictions on α makes clear that Hessian of v should be replaced by Laplacian of v (or some kind of semi-Laplacian-semi-Hessian).

6. EXAMPLES SHOWING ALMOST PERFECT ANALOGY BETWEEN STOCHASTIC OPTIMAL CONTROL AND HARMONIC ANALYSIS

6.1. A_∞ weights and associated Carleson measures. Buckley’s inequality.

We call a nonnegative function on \mathbb{R} an A_∞ weight (dyadic A_∞ weight actually) if

$$(125) \quad \langle w \rangle_J \leq C_1 e^{(\log w)_J}, \forall J \in \mathcal{D}.$$

Here \mathcal{D} is a dyadic lattice on \mathbb{R} , $\langle \cdot \rangle_J$ is the averaging over J . We are going to illustrate our use of Bellman function technique by a collection of examples, the first of which is the result of Buckley that can be found (along with “continuous analogs”) in the paper of Fefferman–Kenig–Pipher [39].

Theorem 6.1. *Let $w \in A_\infty$. Then*

$$(126) \quad \forall I \in \mathcal{D}, \frac{I}{|I|} \sum_{\ell \subseteq I, \ell \in \mathcal{D}} \left(\frac{\langle w \rangle_{\ell_+} - \langle w \rangle_{\ell_-}}{\langle w \rangle_\ell} \right)^2 |\ell| \leq C_2,$$

Where C_2 depends only on C_1 in (125). Here ℓ_\pm are right and left sons of $\ell \in \mathcal{D}$.

Who moves ?

$$x_1, x_2 = \langle w \rangle_J, \langle \log w \rangle_J$$

$$\alpha_1 = \langle w \rangle_{\text{son of } J} - \langle w \rangle_J \Rightarrow |\alpha_1| = \frac{1}{2} |\langle w \rangle_{J_-} - \langle w \rangle_{J_+}|.$$

Function of profit can be read off (126) if one notices that $\frac{1}{|I|} \sum_{\ell \subseteq I, \ell \in \mathcal{D}} \dots$ is the average over the lines of life. Each line of life initiates at I and then proceeds to $I_{\varepsilon_1}(\varepsilon_1 = +1 \text{ or } \varepsilon_1 = -1)$, then to $I_{\varepsilon_1 \varepsilon_2}(\varepsilon_2 = +1 \text{ or } \varepsilon_2 = -1)$, etc.

Thus $\frac{1}{|I|} \sum_{\ell \subseteq I, \ell \in \mathcal{D}} \dots$ plays the role of $\mathbb{E} \int_0^\infty \dots$. This allows us to choose the correct profit function

$$f(\alpha, x) = \frac{4\alpha_1^2}{x_1^2}.$$

Bonus function $F \equiv 0$ here. Bellman equation reads now

$$(127) \quad \sup_{\alpha=(\alpha_1, \alpha_2)} \left[\langle H_v \alpha, \alpha \rangle + \frac{8\alpha_1^2}{x_1^2} \right] = 0$$

to be solved in

$$(128) \quad \Omega = \{(x_1, x_2) : 1 \leq x_1 e^{-x_2} \leq c_1\}$$

with the obstacle condition

$$(129) \quad v(x) \geq 0 \quad \forall x \in \Omega.$$

Compare with (121)!

6.2. A two-weight inequality.

$$\forall J \in \mathcal{D} \quad \langle u \rangle_J \langle v \rangle_J \leq 1 \Rightarrow \forall I \in \mathcal{D}$$

$$\begin{aligned} \frac{1}{|I|} \sum_{\ell \subseteq I, \ell \in \mathcal{D}} |\langle u \rangle_{\ell_+} - \langle u \rangle_\ell| |\langle v \rangle_{\ell_+} - \langle v \rangle_{\ell_-}| |\ell| \\ \leq C \langle u \rangle_I^{1/2} \langle v \rangle_I^{1/2}. \end{aligned}$$

Who moves ?

$$x_1, x_2 = \langle u \rangle_J, \langle v \rangle_J.$$

As in the previous problem $f^\alpha(x)$ is easy to find :

$$f(\alpha, x) = 4|\alpha_1| |\alpha_2|.$$

Bonus function $F \equiv 0$ here again. Bellman equation

$$\sup_{\alpha=(\alpha_1, \alpha_2) \in \mathbb{R}^2} [\langle H_v \alpha, \alpha \rangle + 8|\alpha_1| |\alpha_2|] = 0, v \geq 0 \text{ in } \Omega = \{x = (x_1, x_2) : 0 \leq x_1, x_2; x_1 x_2 \leq 1\}.$$

Compare with (121)!

6.3. John-Nirenberg inequality : Bellman equation with a drift but with $f(\alpha, x) \equiv 0$.

$$\forall J \in \mathcal{D} \langle |\varphi - \langle \varphi \rangle_J|^2 \rangle_J \leq \delta \Rightarrow \forall I \in \mathcal{D}$$

$$\langle e^\varphi \rangle_I \leq C_\delta e^{\langle \varphi \rangle_I}.$$

Who moves ?

$$\begin{aligned} x_1 &= \langle \varphi \rangle_J, \quad x_2 = \langle |\varphi - \langle \varphi \rangle_J|^2 \rangle_J = \\ &= \frac{1}{|J|} \sum_{I \subseteq J, I \in \mathcal{D}} \left\{ \frac{\langle \varphi \rangle_{I_+} - \langle \varphi \rangle_{I_-}}{2} \right\}^2 |I|. \end{aligned}$$

Notice that ($t = n$):

$$x_2^t - \mathbb{E}(x_2^{t+1} | x^t) = x_2 - \frac{x_2^+ + x_2^-}{2} = \left(\frac{x_1^+ - x_1^-}{2} \right)^2 = (\alpha_1^t)^2.$$

But

$$x_1^t - \mathbb{E}(x_1^{t+1} | x^t) = x_1 - \frac{x_1^+ + x_1^-}{2} = 0.$$

On the other hand,

$$x^{t+1} = x^t + \int_t^{t+1} \sigma dw^s + \int_t^{t+1} b ds.$$

Thus drift b stands for $\mathbb{E}(x^{t+1} | x^t) - x^t$ (in the case of discrete time). Therefore, $b(\alpha, x) = \begin{pmatrix} 0 \\ -\alpha_1^2 \end{pmatrix}$ in our case. Notice that $f(\alpha, x) \equiv 0$ as there is no $\frac{1}{|I|} \sum_{\ell \subseteq I} \dots$ in the functional. Bellman equation in this case has a form

$$\sup_{\alpha=(\alpha_1, \alpha_2)} \left[\frac{1}{2} \langle d^2 v \alpha, \alpha \rangle - \frac{\partial v}{\partial x_2} \alpha_1^2 \right] = 0.$$

Compare with (122)!

In other words :

$$(130) \quad \begin{pmatrix} \frac{\partial^2 v}{\partial x_1^2} - 2 \frac{\partial v}{\partial x_2} & \frac{\partial^2 v}{\partial x_1 \partial x_2} \\ \frac{\partial^2 v}{\partial x_1 \partial x_2} & \frac{\partial^2 v}{\partial x_2^2} \end{pmatrix} \leq 0, \det \begin{pmatrix} \frac{\partial^2 v}{\partial x_1^2} - 2 \frac{\partial v}{\partial x_2} & \frac{\partial^2 v}{\partial x_1 \partial x_2} \\ \frac{\partial^2 v}{\partial x_1 \partial x_2} & \frac{\partial^2 v}{\partial x_2^2} \end{pmatrix} = 0.$$

in $\Omega_\delta = \{x = (x_1, x_2), x_1 \in \mathbb{R}, 0 \leq x_2 \leq \delta\}$. The obstacle condition is

$$(131) \quad v(x) \geq F(x) \equiv e^{x_1} \text{ in } \Omega_\delta$$

Denote B_δ^d the dyadic Bellman function of a corresponding problem. There are many solutions of the above equation in Ω_δ which satisfy the obstacle condition $\geq e^{x_1}$ in Ω_δ and even satisfying the boundary condition $= e^{x_1}$ on $x_2 = 0$. These are

$$\varphi_{\varepsilon,q}(x_1, x_2) = q \frac{(1 - \sqrt{\varepsilon - x_2})}{1 - \sqrt{\varepsilon}} e^{x_1 + \sqrt{\varepsilon - x_2} - \sqrt{\varepsilon}}, \delta \leq \varepsilon < 1, q \geq 1.$$

One can compute v_δ -the smallest solution of the above equation satisfying the obstacle condition.

$$v_\delta = \frac{1 - \sqrt{\delta - x_2}}{1 - \sqrt{\delta}} e^{x_1 + \sqrt{\delta - x_2} - \sqrt{\delta}}.$$

This is not B_δ^d ! In fact, $B_\delta^d > v_\delta$. However, v_δ is **the Bellman function for non-dyadic John–Nirenberg inequality!!!** The rest is in Vasyunin’s lectures and in [53].

6.4. Burkholder-Bellman function.

$$\begin{aligned} \forall I \in \mathcal{D} \quad |\langle g \rangle_{I+} - \langle g \rangle_{I-}| &\leq |\langle f \rangle_{I+} - \langle f \rangle_{I-}| \\ \Rightarrow \forall I \in \mathcal{D} \text{ such that } |\langle g \rangle_I| &\leq |\langle f \rangle_I| \end{aligned}$$

one has

$$\langle |g|^p \rangle_I \leq (p-1)^p \langle |f|^p \rangle_I \quad p \geq 2.$$

The constant $(p-1)^p$ is sharp. This is a famous theorem of Burkholder which he proved by constructing the corresponding Bellman function. He found it by solving a corresponding Bellman PDE - a complicated one. We would like to show a simple “heuristic” method of solution.

Who moves ?

$$x_1 = \langle g \rangle_J, \quad x_2 = \langle f \rangle_J, \quad x_3 = \langle |f|^p \rangle_J.$$

Our rules say that $f(\alpha, x) = 0$, $\mathbb{E}F(x_1^t, x_2^t, x_3^t) \approx \mathbb{E}|g|^p$. Denoting by \mathcal{F}_t the σ -algebra generated by dyadic subintervals of I of length $2^{-n}|I|$, $t = 2^n$, we can write $\mathbb{E}|g|^p \approx \mathbb{E}(\mathbb{E}x_1|\mathcal{F}_t)|^p = \mathbb{E}|x_1^t|^p$ which gives us the correct bonus function $F(x_1, x_2, x_3) = |x_1|^p$. Notice that $A = \{\alpha = (\alpha_1, \alpha_2, \alpha_3) : |\alpha_1| \leq |\alpha_2|\}$ now.

This is because $|\alpha_1| = \frac{1}{2}|\langle g \rangle_{J+} - \langle g \rangle_{J-}|$, $|\alpha_2| = \frac{1}{2}|\langle f \rangle_{J+} - \langle f \rangle_{J-}|$, and we are given that the first quantity is always majorized by the second one.

So we have the Bellman equation

$$\sup_{|\alpha_1| \leq |\alpha_2|, \alpha_3} \langle H_v \alpha, \alpha \rangle = 0$$

in $\Omega = \{x : (x_1, x_2, x_3) : |x_2|^p \leq x_3\}$ (convex), with obstacle condition

$$v(x_1, x_2, x_3) \geq |x_1|^p.$$

Compare with (121)!

This example is interesting because we have a non-trivial set of restrictions A for “control” α .

Solutions were given by Burkholder [19] (see also [20]–[25]) and also (a different approach using Monge–Ampère equation) can be found in [59]. See also a very interesting review [4].

An interesting Bellman function built by the use of Monge–Ampère equation can be also found in [60], [61].

7. THE TECHNIQUE OF LAMINATES, BELLMAN FUNCTION, AND ESTIMATES OF SINGULAR INTEGRALS FROM BELOW

Definition. Laminate on $M_{2 \times 2}^s$ is a positive finite measure on symmetric real matrices $M_{2 \times 2}^s$ such that

$$(132) \quad f(A) \geq \int f(A + M) d\nu(M)$$

for all rank 1 concave functions f .

Theorem 7.1. Any laminate on $M_{2 \times 2}^s$ can be approximated weakly by the push forward of Lebesgue measure on the plane by the Hessian of smooth compactly supported functions, in other words, for **any** good F

$$\int F(M) d\nu(M) \approx \int_{\mathbb{R}^2} F(Du) dx dy,$$

$$\text{where } Du(x, y) := \begin{bmatrix} u_{xx} & u_{xy} \\ u_{yx} & u_{yy} \end{bmatrix}.$$

Observation. Laminates supported by diagonal matrices $\begin{bmatrix} X & 0 \\ 0 & Y \end{bmatrix}$ are just exactly exactly the measures on \mathbb{R}^2 such that $(z = (X, Y))$

$$(133) \quad f(a) \geq \int_{\mathbb{C}} f(a + z) d\nu(z)$$

for all **bi-concave** (meaning separately concave in X and Y) function f .

Definition. $(\int X d\nu, \int Y d\nu)$ is called baricenter of a laminate ν supported on diagonal matrices.

Fix $p > 2$ and $p_\eta = p + \eta, \eta > 0$. Put

$$s_0 := 1 - \frac{2}{p_\eta}, K := \frac{p_\eta}{p_\eta - 2}, p - \eta - 1 = \frac{K + 1}{K - 1}.$$

We are going to construct very interesting laminates supported on

$$Y = KX, Y + \frac{1}{K}X.$$

Fix $p \geq 2$, fix small $\eta > 0$, put

$$p_\eta := p + \eta,$$

$$(134) \quad s_0 := 1 - \frac{2}{p_\eta}, K := \frac{1}{s_0} = \frac{p_\eta}{p_\eta - 2}, p_\eta = \frac{2K}{K - 1}, p_\eta - 1 = \frac{K + 1}{K - 1}$$

We are going to present an interesting laminate with baricenter $(1, 1)$ supported by lines

$$L_K : Y = KX, L_{1/K} : Y = \frac{1}{K}X.$$

Let f be a bi-concave function and

$$(135) \quad f(z) = O(|z|^p), \quad z \rightarrow \infty.$$

Then concavity in horizontal variable gives

$$(136) \quad f(t, t+h) \geq \frac{t - \frac{1}{K}(t+h)}{t+h - \frac{1}{K}(t+h)} f(t+h, t+h) + \frac{h}{t+h - \frac{1}{K}(t+h)} f\left(\frac{1}{K}(t+h), t+h\right).$$

Rewrite it as

$$(137) \quad f(t+h, t+h) \leq \frac{t+h - \frac{1}{K}(t+h)}{t - \frac{1}{K}(t+h)} f(t, t+h) - \frac{h}{t - \frac{1}{K}(t+h)} f\left(\frac{1}{K}(t+h), t+h\right).$$

The concavity in vertical variable gives

$$(138) \quad f(t, t) \geq \frac{t - \frac{1}{K}t}{t - \frac{1}{K}t + h} f(t, t+h) + \frac{h}{t - \frac{1}{K}t + h} f\left(t, \frac{1}{K}t\right).$$

From (137), (138) we obtain (of course we divide by h , and next, we will make h tend to 0)

$$\begin{aligned} \frac{f(t+h, t+h) - f(t, t)}{h} &\leq \frac{1}{h} \left[\frac{t+h - \frac{1}{K}(t+h)}{t - \frac{1}{K}(t+h)} - 1 + 1 - \frac{t - \frac{1}{K}t}{t - \frac{1}{K}t + h} \right] f(t, t+h) - \\ &\quad \frac{1}{t - \frac{1}{K}(t+h)} f\left(\frac{1}{K}(t+h), t+h\right) - \frac{1}{t - \frac{1}{K}t + h} f\left(t, \frac{1}{K}t\right). \end{aligned}$$

Make $h \rightarrow 0$. Then

$$(139) \quad f'(t, t) - \frac{2K}{K-1} \frac{f(t, t)}{t} \leq -\frac{K}{K-1} \frac{f(\frac{1}{K}t, t)}{t} - \frac{K}{K-1} \frac{f(t, \frac{1}{K}t)}{t}.$$

We recall (134) and multiply by $1/t^{p\eta}$. Notice that after that $LHS = \left(\frac{f(t, t)}{t^{p\eta}} \right)'$.

We integrate from 1 to ∞ and use (135) to forget the term at infinity. Then we obtain for any bi-concave function on the plane

$$(140) \quad -f(1, 1) \leq -\frac{K}{K-1} \int_1^\infty f\left(\frac{1}{K}t, t\right) \frac{dt}{t^{p\eta+1}} - \frac{K}{K-1} \int_1^\infty f\left(t, \frac{1}{K}t\right) \frac{dt}{t^{p\eta+1}}$$

Introduce $\nu_{K, \eta}$:

$$\int_{\mathbb{R}^2} \phi d\nu_{K, \eta} = \frac{K}{K-1} \int_1^\infty \phi\left(\frac{1}{K}t, t\right) \frac{dt}{t^{p\eta+1}}.$$

It is a laminate supported by $L_K : Y = KX$. And introduce $\nu_{1/K, \eta}$:

$$\int_{\mathbb{R}^2} \phi d\nu_{1/K, \eta} = \frac{K}{K-1} \int_1^\infty \phi\left(t, \frac{1}{K}t\right) \frac{dt}{t^{p\eta+1}}.$$

It is a laminate supported by $L_{1/K} : Y = \frac{1}{K}X$. Now (140) can be rewritten as

$$(141) \quad f(1, 1) \geq \int f(d\nu_{K, \eta} + d\nu_{1/K, \eta})$$

If all our concavity in getting (141) become linearities then we have equality in (141). So $d\nu_{K,\eta} + d\nu_{1/K,\eta}$ is a laminate with baricenter $(1, 1)$.

Consider a new laminate, now with baricenter $(0, 0)$:

$$\mu_{K,\eta} = \frac{1}{4}(d\nu_{K,\eta} + d\nu_{1/K,\eta}) + \frac{1}{4}\delta_{(-1,1)} + \frac{1}{2}\delta_{(0,1)}.$$

Test it on

$$\phi_1(X, Y) = |X + Y|^p, \quad \phi_2(X, Y) = |X - Y|^p.$$

Then

$$(142) \quad \frac{\int \phi_1 d\mu_{K,\eta}}{\int \phi_2 d\mu_{K,\eta}} = \frac{\frac{1}{4}K((K+1)^p + (K+1)^p/K^p)\eta^{-1} + \frac{1}{2}(K-1)}{\frac{1}{4}K((K-1)^p + (K-1)^p/K^p)\eta^{-1} + \frac{1}{2}(K-1) + \frac{1}{4}2^p(K-1)}.$$

Choosing $\eta > 0$ very small we get

$$(143) \quad \frac{\int \phi_1 d\mu_{K,\eta}}{\int \phi_2 d\mu_{K,\eta}} \geq \left(\frac{K+1}{K-1}\right)^p - C\eta.$$

Notice that we can consider a bit different laminate than $\mu_{K,\eta}$, Namely let us push forward $\mu_{K,\eta}$ by the map $X \rightarrow X, Y \rightarrow -Y$. The new measure is called $\sigma_{K,\eta}$. Then (143) transforms to

$$(144) \quad \frac{\int \phi_2 d\sigma_{K,\eta}}{\int \phi_1 d\sigma_{K,\eta}} \geq \left(\frac{K+1}{K-1}\right)^p - C\eta.$$

Now we use Theorem 7.1. It implies that there exist smooth functions with compact support on the plane such that

$$(145) \quad \frac{\int |u_{xx} - u_{yy}|^p dm_2}{\int |u_{xx} + u_{yy}|^p dm_2} \geq \left(\frac{K+1}{K-1}\right)^p - C\eta.$$

Notice that K depends on η (see (134)) but

$$\frac{K+1}{K-1} \rightarrow p-1, \quad \eta \rightarrow 0.$$

Thus from (145) we get the estimate

$$(146) \quad \|R_1^2 - R_2^2\|_p \geq p-1,$$

if $p \geq 2$.

This argument can be applied to some other interesting singular operators. Constant $p^* - 1$ can be described as the **smallest** constant $c = c_p$ such that the function

$$h_c(X, Y) = |Y + X|^p - c^p |Y - X|^p$$

has a bi-concave majorant.

Definition. Let us call $\varphi_p(X, Y)$ the smallest bi-concave majorant of $h_c(X, Y) = |Y + X|^p - c^p |Y - X|^p$ for the smallest (as we know) possible $c = c_p = p^* - 1$.

We will recall a formula for φ_p in the next section.

Now let us consider a different family (it is a perturbation of h_c):

$$h_{c,\tau} := |((Y + X)^2 + \tau^2(X - Y)^2)^{1/2}|^p - c^p |Y - X|^p.$$

Here is a result proved in [18].

Theorem 7.2. *For sufficiently small universal $\tau_0 > 0$, any $p \in (1, \infty)$, and any $\tau \in [-\tau_0, \tau_0]$, the smallest c for which there exists a bi-concave majorant of $h_{c,\tau}$ is $c_p(\tau) = ((p^* - 1)^2 + \tau^2)^{1/2}$.*

Using the same considerations with laminates as above (especially Theorem 7.1) we can prove the following estimate from below for “quantum” linear combination of second order Riesz transforms:

Theorem 7.3. *For sufficiently small τ and any small positive ϵ one can find $g \in L^p(m_2)$ such that*

$$\|(|(R_1^2 - R_2^2)g|^2 + \tau^2|(R_1^2 + R_2^2)g|^2)^{1/2}\|_p \geq ((p^* - 1)^2 + \tau^2)^{1/2}\|g\|_p - \epsilon.$$

This gives rise to the following problem:

Problem. For sufficiently small τ

$$\left\| \begin{bmatrix} R_1^2 - R_2^2 \\ \tau I \end{bmatrix} : L^p(m_2) \rightarrow L^p(\mathbb{R}^2, l^2) \right\| = ((p^* - 1)^2 + \tau^2)^{1/2} ?$$

The answer is affirmative, see [18]. Notice that for $p \in (1, 2)$ and large τ this is no longer true. Somewhere we have a “phase transition” of the sharp constant. It is not clear what is the critical $\tau(p)$.

7.1. “Explanation” of laminates above via Burkholder’s function $\varphi_p(X, Y)$ and its properties. We introduce coordinates (x, y) :

$$Y = y + x, \quad X = y - x.$$

Let

$$\gamma_p = p(1 - \frac{1}{p^*})^{p-1}.$$

In the first and second quadrants of xy , Burkholder’s function in these coordinates is equal to (here the reader should glance at (134) and make $\eta = 0$ in it, s_0 and k below are as in (134), but with $\eta = 0$)

$$(147) \quad \varphi_p(x, y) := \begin{cases} \gamma_p(y - (p^* - 1)|x|)(|x| + y)^{p-1}, & \text{if } \frac{y-|x|}{y+|x|} \geq s_0 := 1 - \frac{2}{p} = \frac{1}{k} \\ y^p - (p^* - 1)^p |x|^p, & \text{if } -1 \leq \frac{y-|x|}{y+|x|} \leq s_0. \end{cases}$$

Now **extend** $\varphi_p(x, y)$ to the whole plane by

$$\varphi_p(x, y) = \varphi_p(-x, -y).$$

Burkholder proved [19]

Theorem 7.4. *Such a function coincides with the smallest majorant of $h_c(x, y) = |y|^p - c^p|x|^p$, $c = p^* - 1$ bi-convex in X, Y coordinates. For $c \in [0, p^* - 1)$ there is no such bi-concave majorant of h_c .*

Observation 2. We use here both coordinates (X, Y) and (x, y) . In the cone $X \leq Y \leq KX$ function φ_p is linear along $Y = \text{const}$ segments. Similarly, In the cone $\frac{1}{K}X \leq Y \leq X$ function φ_p is linear along $X = \text{const}$ segments.

This linearity allows to calculate (we are in (X, Y) now)

$$\varphi_p(t + h, t + h) - \varphi_p(t, t)$$

virtually without any loss if we use the T -shape 4-tuple of points in \mathbb{R}^2 : $((\frac{1}{K}(t+h), t+h), (t, t+h), (t+h, t+h), (t, \frac{1}{K}t))$ as in Section 7.

If we move one of the lines $L_K, L_{1/K}$ then two things may happen: 1) we go outside of these linearity cones, and subsequently we get strict inequality for $\varphi_p(1, 1)$, or 2) if we do not go outside of linearity cones, but then we loose

$$\varphi_p = h_{p^*-1}$$

equality because by the definition of φ_p (see (147)) this equality holds only on the boundary of and *outside* of the union of linearity cones.

Notice also that on these lines $L_K, L_{1/K}$ (recall that $K = \frac{p+\eta}{p+\eta-2}$ if $p \geq 2$) we have that

$$\varphi_p \approx h_{p^*-1} \approx 0, \varphi_p \geq h_{p^*-1}.$$

For c larger than $p^* - 1$ we can again choose the lines where h_c coincides with its bi-convex majorant, but then they will be quite negative there and integration of h_c along a laminate supported on such lines cannot be almost positive as it was the case above.

8. STOCHASTIC CALCULUS AND 1/2 QUASICONVEXITY

We have a bijection of matrices $M = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ onto (z, w) : $z = a + d + i(b - c)$, $w = a - d + i(b + c)$. Notice that $2 \det M = |z|^2 - |w|^2$.

Recall that Sverak's function is the following "simple" object

$$S(z, w) := \begin{cases} |z|^2 - |w|^2, & |z| + |w| \leq 1 \\ 2|z| - 1, & \text{otherwise} \end{cases}$$

Function

$$\psi_p(z, w) := ((p-1)|z| - |w|)(|z| + |w|)^{p-1}, p \geq 2,$$

can be easily obtained from S using the idea of Iwaniec, see e.g. [9]. The process is a certain averaging. Therefore, the fact that S is rank-1 convex implies that ψ_p is also rank-1 convex.

To solve the Big Iwaniec problem of the previous sections it would be enough that any of these functions is quasiconvex at zero matrix. This is an outstanding and very difficult problem.

On the other hand we can formulate two problems which seem to be easier and may be readily reachable by Stochastic Calculus methods:

Problem. Prove that $S(z, \frac{1}{2}w)$ is a quasiconvex function at zero matrix.

At least we feel that the following problem is directly reachable by methods of Stochastic Calculus:

Problem. Prove that $\psi_p(z, \frac{1}{2}w), p \geq 2$, is a quasiconvex function at zero matrix.

See interesting results in recent paper [3].

REFERENCES

- [1] Kari Astala, *Area distortion of quasiconformal mappings*, Acta Math., v. 173 (1994), 37–60.
- [2] Kari Astala, Tadeusz Iwaniec, Eero Saksman, *Beltrami operators on the plane*, Duke Math. J., v. 107 (2001), pp. 27–56.

- [3] Kari Astala, Tadeusz Iwaniec, Istvn Prause, Eero Saksman, *Burkholder integrals, Morrey's problem and quasiconformal mappings*, arXiv:1012.0504 .
- [4] R. Bañuelos, *The foundational inequalities of D. L. Burkholder and some of their ramifications*, Illinois J. of Math., volume in honor of D. L. Burkholder, to appear.
- [5] R. Bañuelos, P. Janakiraman, *L^p -bounds for the Beurling–Ahlfors transform*, Trans. Amer. Math. Soc., **360**, (2008), no. 7, 3603–3612.
- [6] R. Banuelos, A. Lindeman, *A martingale study of the Beurling–Ahlfors transform in \mathbb{R}^n* , J. Funct. Analysis, **145**, (1997), 224–265.
- [7] R. Banuelos, P. J. Méndez-Hernandez, *Space-time Brownian motion and the Beurling–Ahlfors transform*, Indiana Univ. Math. J., **52**, (2003), no. 4, 981–990.
- [8] R. Bañuelos, G. Wang, *Sharp inequalities for martingales with applications to the Beurling–Ahlfors and Riesz transforms*, Duke Math. J., **80**, No. 3, (1995), 575–600.
- [9] A. Baernstein, S. J. Montgomery-Smith, *Some conjectures about integral means of ∂f and $\bar{\partial} f$* , in “Complex Analysis and Differential Equations”, Proc. of Marcus Wallenberg Symposium in honor of Matts Essen, ed. Ch. Kiselman, Uppsala, Sweden 1999, 92–109.
- [10] B. V. Bojarski, *Homeomorphic solutions of Beltrami systems*, Dokl. Akad. Nauk. SSSR, **102**, (1955), 661–664.
- [11] B. V. Bojarski, *Generalized solutions of a system of differential equations of first order and elliptic type with discontinuous coefficients*, Mat. Sbornik, No 43, **85**, (1957), 451–503.
- [12] B. V. Bojarski, *Quasiconformal mappings and general structure properties of systems of nonlinear equations elliptic in the sense of Lavrentiev*, Symposia Mathematica (1976).
- [13] B. V. Bojarski T. Iwaniec, *Quasiconformal mappings and non-linear elliptic equations in two variables I,II*, Bull. Acad. Pol. Sci., **22**, No. 5, (1974), 473–484.
- [14] J. BENNETT, N. BEZ, A. CARBERY, D. HUNDERTMARK *Heat flow and Strichartz*. Personal communication.
- [15] J. BENNETT, A. CARBERY, T. TAO *On the multilinear restriction and Kakeya conjectures*, arxiv: math/0509262v1 12 Sep 2005.
- [16] A. BORICHEV, P. JANAKIRAMAN, A. VOLBERG, *On burkholder function for ortogonal martingales and zeros of Legendre polynomials*, arXiv.
- [17] A. BORICHEV, P. JANAKIRAMAN, A. VOLBERG, *Subordination by orthogonal martingales in L^p and zeros of Laguerre polynomials*. Preprint, pp. 1–19, 2010.
- [18] N. BOROS, P. JANAKIRAMAN, A. VOLBERG, *Burkholder function for “quantum” linear combination of Riesz transforms*. Preprint, pp. 1–28, 2010.
- [19] D. BURKHOLDER, *Boundary value problems and sharp estimates for the martingale transforms*, Ann. of Prob. **12** (1984), 647–702.
- [20] D. BURKHOLDER, *An extension of classical martingale inequality*, Probability Theory and Harmonic Analysis, ed. by J.-A. Chao and W. A. Woyczynski, Marcel Dekker, 1986.
- [21] D. BURKHOLDER, *Sharp inequalities for martingales and stochastic integrals*, Colloque Paul Lévy sur les Processus Stochastiques (Palaiseau, 1987), Astérisque No. 157–158 (1988), 75–94.
- [22] D. BURKHOLDER, *Differential subordination of harmonic functions and martingales*, (El Escorial 1987) Lecture Notes in Math., **1384** (1989), 1–23.
- [23] D. BURKHOLDER, *Explorations of martingale theory and its applications*, Lecture Notes in Math. **1464** (1991), 1–66.
- [24] D. BURKHOLDER, *Strong differential subordination and stochastic integration*, Ann. of Prob. **22** (1994), 995–1025.
- [25] D. BURKHOLDER, *A proof of the Peczynski’s conjecture for the Haar system*, Studia MATH., **91** (1988), 79–83.
- [26] Michael Christ, *A $T(b)$ theorem with remarks on analytic capacity and the Cauchy integral*, Colloq. Math. **60/61** (1990), no. 2, 601–628.
- [27] D. Cruz-Uribe, J. M. Martell, C. Prez, “Weights, Extrapolation and the Theory of Rubio de Francia”, Operator Theory: Advances and Applications, Vol. 215, Springer, 2011.
- [28] B. DACOROGNA, *Some recent results on polyconvex, quasiconvex and rank one convex functions*, Adv. Math. Appl. Sci., World Science Publ., 1994, pp. 169–176.
- [29] B. DACOROGNA, *Direct Methods in the Calculus of Variations*, Springer, 1989.

- [30] O. DRAGICEVIC, A. VOLBERG *Sharp estimates of the Ahlfors-Beurling operator via averaging of Martingale transform*, Michigan Math. J. **51** (2003), 415-435.
- [31] O. DRAGICEVIC, A. VOLBERG, *Bellman function, Littlewood-Paley estimates, and asymptotics of the Ahlfors-Beurling operator in $L^p(\mathbb{C})$, $p \rightarrow \infty$* , Indiana Univ. Math. J. **54** (2005), no. 4, 971-995.
- [32] O. DRAGICEVIC, A. VOLBERG, *Bellman function and dimensionless estimates of classical and Ornstein-Uhlenbeck Riesz transforms*. J. of Oper. Theory, **56** (2006) No. 1, pp. 167-198.
- [33] O. DRAGICEVIC, S. PETERMICHL, A. VOLBERG, *A rotation method which gives linear L^p -estimates for powers of the Ahlfors-Beurling operator*. Journal des Mathématiques Pures et Appliquées, **86**, No. 6 (2006), 492-509.
- [34] O. DRAGICEVIC, S. TREIL, A. VOLBERG, *A lemma about 3 quadratic forms*, arXiv:0710.3249. To appear in Intern. Math. Research Notices.
- [35] S. GEISS, S. MONTGOMERY-SMITH, E. SAKSMAN, *On singular integral and martingale transforms*, arxiv: math. CA/0701516v1 18 June 2007.
- [36] S. Buckley *Estimates for operator norms on weighted spaces and reverse Jensen inequalities*, Trans. Amer. Math. Soc., **340**, (1993), 253-273.
- [37] Javier Duoandikoetxea, *Extrapolation of weights revisited: new proofs and sharp bounds*. J. Funct. Anal., v. 260 (2011), 1886-1901.
- [38] Oliver Dragicevic, Loukas Grafakos, Cristina Pereyra, Stefanie Petermichl, *Extrapolation and sharp norm estimates for classical operators on weighted Lebesgue spaces*, Publ. Mat., v. 49 ((2005), 73-91.
- [39] R. Fefferman, C. Kenig, J. Pipher *The theory of weights and the Dirichlet problem for elliptic equations*, Annals of Math., **134**, (1991), 65-124.
- [40] J. Garcia-Cuerva, J. Rubio de Francia *Weighted Norm Inequalities And Related Topics*, North-Holland Mathematics Studies, 116, North-Holland, Amsterdam-New York-Oxford, 1985.
- [41] F. W. Gehring, *Open problems*, Proceedings of Rumanian-Finnish Seminar on Teichmüller Spaces and Quasiconformal Mappings, 1969, page 306.
- [42] F. W. Gehring, *The L^p -integrability of the partial derivatives of a quasiconformal mapping*, Acta Math., **130** (1973), 265-277.
- [43] F. W. Gehring, *Topics in quasiconformal mappings*, Proceedings of the ICM 1986, Berkeley, 62-80.
- [44] F. W. Gehring, E. Reich *Area distortion under quasiconformal mappings*, Ann. Acad. Sci. Fenn. Ser. AI, **388**, (1966), 1-15.
- [45] I. I. Gihkman, A. V. Skhorohod *The theory of Stochastic Processes, I, II*, Springer-Verlag, 1974.
- [46] T. Iwaniec *Extremal inequalities in Sobolev spaces and quasiconformal mappings*, Z. Anal. Anwendungen, **1**, (1982), 1-16.
- [47] N. Krylov, *Optimal Control of Stochastic Processes*. Springer-Verlag, 1980.
- [48] S. Konyagin, A. Volberg, *On measures with the doubling condition*. Izv. Akad. Nauk SSSR Ser. Mat. **51** (1987), no. 3, 666-675; translation in Math. USSR-Izv. **30** (1988), no. 3, 629-638.
- [49] F. Nazarov, A. Volberg, *Heat extension of the Ahlfors-Beurling operator and estimates of its norms*, Algebra i Analiz, **15**, (2003), no. 4, 142-158, translated in St. Petersburg Math. J.
- [50] Nazarov, F.; Treil, S.; Volberg, A. *The Bellman functions and two-weight inequalities for Haar multipliers*. J. Amer. Math. Soc. **12** (1999), no. 4, 909-928.
- [51] S. Petermichl, L. Slavin, B. D. Wick, *New estimates for the Beurling-Ahlfors operator on differential forms*. arXiv:0901.0345
- [52] L. Slavin, A. Volberg, *Bellman Function and the $H^1 - BMO$ Duality* arXiv:0809.0322, Journal-ref: Contemporary Math, 428, AMS, 2007
- [53] L. Slavin, V. Vasyunin, *Sharp results in the integral-form John-Nirenberg inequality*. arXiv:0709.4332
- [54] A. Stokolos, L. Slavin, V. Vasyunin, *Bellman function for maximal operator*.
- [55] E. Stein, *Singular Integrals and Differentiability Properties of Functions*.

- [56] S. Petermichl, J. Wittwer, A sharp estimate for the weighted Hilbert transform via Bellman functions. *Michigan Math. J.* 50 (2002), no. 1, 7187.
- [57] Petermichl, S.; Wittwer, J. Heating of the Beurling operator: sufficient conditions for the two-weight case. *Studia Math.* 186 (2008), no. 3, 203217.
- [58] S. Petermichl, A. Volberg, Heating of the Ahlfors–Beurling operator: weakly quasiregular maps on the plane are quasiregular. *Duke Math. J.*, v. 112 (2002), 281–305.
- [59] V. Vasyunin, A. Volberg, Burkholder’s function via Monge–Ampère equation, arXiv:1006.2633.
- [60] V. Vasyunin, A. Volberg, Monge–Ampère equation and Bellman optimization of Carleson embedding theorems. *Linear and complex analysis*, 195238, Amer. Math. Soc. Transl. Ser. 2, 226, Amer. Math. Soc., Providence, RI, 2009.
- [61] V. Vasyunin, A. Volberg, Notes on Bellman functions in Harmonic Analysis, at blog sashavolberg.wordpress.com
- [62] A. D. Wentzel *Introduction to Stochastic Integrals*, Moscow, Nauka, 1972.
- [63] J. Wittwer, A sharp estimate on the norm of the martingale transform. *Math. Res. Lett.* 7 (2000), no. 1, 112.

DEPARTMENT OF MATHEMATICS, MICHIGAN STATE UNIVERSITY, EAST LANSING, MI 48824,
USA

E-mail address: sashavolberg@yahoo.com